

Série Investigação Filosófica

Textos selecionados de

Filosofia da Economia

André Nascimento Pontes
Mariana Kuhn de Oliveira
Ramiro de Ávila Peres
(Organizadores)

TEXTOS SELECIONADOS DE FILOSOFIA DA ECONOMIA

Série Investigação Filosófica

TEXTOS SELECIONADOS DE FILOSOFIA DA ECONOMIA

André Nascimento Pontes
Mariana Kuhn de Oliveira
Ramiro de Ávila Peres
(Organizadores)



Pelotas, 2022

REITORIA

Reitora: Isabela Fernandes Andrade

Vice-Reitora: Ursula Rosa da Silva

Chefe de Gabinete: Aline Ribeiro Paliga

Pró-Reitor de Graduação: Maria de Fátima Cossio

Pró-Reitor de Pesquisa e Pós-Graduação: Flávio Fernando Demarco

Pró-Reitor de Extensão e Cultura: Eraldo dos Santos Pinheiro

Pró-Reitor de Planejamento e Desenvolvimento: Paulo Roberto Ferreira Júnior

Pró-Reitor Administrativo: Ricardo Hartlebem Peter

Pró-Reitor de Gestão de Informação e Comunicação: Julio Carlos Balzano de Mattos

Pró-Reitor de Assuntos Estudantis: Fabiane Tejada da Silveira

Pró-Reitor de Gestão Pessoas: Taís Ulrich Fonseca

CONSELHO EDITORIAL DA EDITORA DA UFPEL

Presidente do Conselho Editorial: Ana da Rosa Bandeira

Representantes das Ciências Agrônômicas: Victor Fernando Büttow Roll

Representantes da Área das Ciências Exatas e da Terra: Eder João Lenardão

Representantes da Área das Ciências Biológicas: Rosângela Ferreira Rodrigues

Representante da Área das Engenharias e Computação: Reginaldo da Nóbrega Tavares

Representantes da Área das Ciências da Saúde: Fernanda Capella Rugno

Representante da Área das Ciências Sociais Aplicadas: Daniel Lena Marchiori Neto

Representante da Área das Ciências Humanas: Charles Pereira Pennaforte

Representantes da Área das Linguagens e Artes: Lúcia Bergamaschi Costa Weymar

EDITORIA DA UFPEL

Chefia: Ana da Rosa Bandeira (Editora-chefe)

Seção de Pré-produção: Isabel Cochrane (Administrativo)

Seção de Produção: Suelen Aires Böettge (Administrativo)

Anelise Heidrich (Revisão)

Ingrid Fabiola Gonçalves (Diagramação)

Seção de Pós-produção: Madelon Schimmelpfennig Lopes (Administrativo)

Morgana Riva (Assessoria)

CONSELHO EDITORIAL

Prof. Dr. João Hobuss (Editor-Chefe)
Prof. Dr. Juliano Santos do Carmo (Editor-Chefe)
Prof. Dr. Alexandre Meyer Luz (UFSC)
Prof. Dr. Rogério Saucedo (UFSM)
Prof. Dr. Renato Duarte Fonseca (UFSM)
Prof. Dr. Arturo Fatturi (UFFS)
Prof. Dr. Jonadas Techio (UFRGS)
Profa. Dra. Sofia Alborno Stein (UNISINOS)
Prof. Dr. Alfredo Santiago Culleton (UNISINOS)
Prof. Dr. Roberto Hofmeister Pich (PUCRS)
Prof. Dr. Manoel Vasconcellos (UFPEL)
Prof. Dr. Marco Antônio Caron Ruffino (UNICAMP)
Prof. Dr. Evandro Barbosa (UFPEL)
Prof. Dr. Ramón del Castillo (UNED/Espanha)
Prof. Dr. Ricardo Navia (UDELAR/Uruguai)
Profa. Dra. Mónica Herrera Noguera (UDELAR/Uruguai)
Profa. Dra. Mirian Donat (UEL)
Prof. Dr. Giuseppe Lorini (UNICA/Itália)
Prof. Dr. Massimo Dell'Utri (UNISA/Itália)

COMISSÃO TÉCNICA (EDITORAÇÃO)

Prof. Dr. Juliano Santos do Carmo (Diagramador/Capista)

DIREÇÃO DO IFISP

Prof. Dr. João Hobuss

CHEFE DO DEPARTAMENTO DE FILOSOFIA

Prof. Dr. Juliano Santos do Carmo

Série Investigação Filosófica

A Série Investigação Filosófica, uma iniciativa do **Núcleo de Ensino e Pesquisa em Filosofia** do Departamento de Filosofia da UFPel e do **Grupo de Pesquisa Investigação Filosófica** do Departamento de Filosofia da UNIFAP, sob o selo editorial do NEPFil online e da Editora da Universidade Federal de Pelotas, tem por objetivo precípuo a publicação da tradução para a língua portuguesa de textos selecionados a partir de diversas plataformas internacionalmente reconhecidas, tal como a *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/>), por exemplo. O objetivo geral da série é disponibilizar materiais bibliográficos relevantes tanto para a utilização enquanto material didático quanto para a própria investigação filosófica.

EDITORES DA SÉRIE

Rodrigo Reis Lastra Cid (IF/UNIFAP)

Juliano Santos do Carmo (NEPFIL/UFPEL)

COMISSÃO TÉCNICA

Marco Aurélio Scarpino Rodrigues (Revisor em Língua Portuguesa)

Rafaela Nobrega (Diagramadora/Capista)

ORGANIZADORES DO VOLUME

André Nascimento Pontes (UFAM)

Mariana Kuhn de Oliveira (CEBRAP)

Ramiro de Ávila Peres (UFRGS)

TRADUTORES E REVISORES

André Nascimento Pontes (UFAM)

Mariana Kuhn de Oliveira (CEBRAP)

Ramiro de Ávila Peres (UFRGS)

Bruno Aislã Gonçalves dos Santos (UNICENTRO)

Thomas Hyeono Kang (FGV EESP)

CRÉDITO DA IMAGEM DE CAPA

BEUCKELAER, Joachim. The Vegetable Market. Source: <https://www.wikidata.org/wiki/>



GRUPO DE PESQUISA INVESTIGAÇÃO FILOSÓFICA (UNIFAP/CNPq)

O Grupo de Pesquisa Investigação Filosófica (DPG/CNPq) foi constituído por pesquisadores que se interessam pela investigação filosófica nas mais diversas áreas de interesse filosófico. O grupo foi fundado em 2010, como grupo independente, e se oficializou como grupo de pesquisa da Universidade Federal do Amapá em 2019.

MEMBROS PERMANENTES DO GRUPO

Aluizio de Araújo Couto Júnior
Bruno Aislã Gonçalves dos Santos
Cesar Augusto Mathias de Alencar
Daniel Schiochett
Daniela Moura Soares
Everton Miguel Puhl Maciel
Guilherme da Costa Assunção Cecílio
Kherian Galvão Cesar Gracher
Luiz Helvécio Marques Segundo
Paulo Roberto Moraes de Mendonça
Pedro Merlussi
Rafael César Pitt
Rafael Martins
Renata Ramos da Silva
Rodrigo Alexandre de Figueiredo
Rodrigo Reis Lastra Cid
Sagid Salles
Tiago Luís Teixeira de Oliveira

© Série Investigação Filosófica, 2022

Universidade Federal de Pelotas
Departamento de Filosofia
Núcleo de Ensino e Pesquisa em Filosofia
Editora da Universidade Federal de Pelotas

Universidade Federal do Amapá
Departamento de Filosofia
Grupo de Pesquisa Investigação Filosófica

NEPFil online

Rua Alberto Rosa, 154 – CEP 96010-770 – Pelotas/RS

Os direitos autorais estão de acordo com a Política Editorial do NEPFil online. As revisões ortográficas e gramaticais foram realizadas pelos tradutores e revisores. A autorização para a tradução dos verbetes da *Stanford Encyclopedia of Philosophy* neste volume foi obtida pelo *Grupo de Pesquisa Investigação Filosófica*.

Primeira publicação em 2022 por NEPFil online e Editora da UFPel.

Dados Internacionais de Catalogação

N123 Textos selecionados de filosofia da economia.
[recurso eletrônico] Organizadores: André Nascimento Pontes, Mariana Kuhn de Oliveira, Ramiro de Ávila Peres – Pelotas: NEPFIL Online, 2022.
319p. - (Série Investigação Filosófica).
Modo de acesso: Internet
<wp.ufpel.edu.br/nepfil>
ISBN: 978-65-86440-83-7

1. Filosofia. 2. Economia. I. Pontes, André Nascimento. II. Oliveira, Mariana Kuhn. III. Peres, Ramiro de Ávila.

COD 100



Para maiores informações, visite o site wp.ufpel.edu.br/nepfil

SUMÁRIO

Sobre a série Investigação Filosófica	13
Introdução	15
(I) Filosofia da Economia	19
1. Introdução: O que é Economia?	20
1.1 O surgimento da ciência econômica e das economias	20
1.2 A economia contemporânea e suas diversas escolas	23
2. Seis problemas metodológicos centrais	26
2.1 Economia positiva versus normativa	26
2.2 Razões versus causas	28
2.3 Naturalismo científico social	29
2.4 Abstração, idealização e cláusulas ceteris paribus em economia	30
2.5 Causas em economia e econometria	31
2.6 Estrutura e estratégia da economia: paradigmas e programas de pesquisa	33
3. Inexatidão, cláusulas ceteris paribus, tendências, “pressupostos irrealistas” e modelos	34
3.1 A economia clássica e o método a priori	34
3.2 Friedman e a defesa dos “pressupostos irrealistas”	38
3.3 Modelos	41
4. Abordagens influentes à metodologia econômica	42
4.1 Abordagens popperianas	42
4.2 A retórica da economia	46
4.3 “Realismo” na metodologia econômica	47
4.4 Metodologia econômica e estudos sociais da ciência	48
4.5 Estudos de caso	50
5. Teoria da escolha racional	54
5.1 Racionalidade individual	54
5.2 Racionalidade coletiva e escolha social	57

5.3 Teoria dos jogos	59
6. Economia e ética	59
6.1 Bem-estar	60
6.2 Eficiência	61
6.3 Outras direções na economia normativa	65
7. Conclusões	66
Referência bibliográfica	66

(II) Teoria da Decisão **119**

1. O que são preferências sobre prospectos?	120
2. Medidas de utilidade para as preferências	124
2.1 Utilidade ordinal	124
2.2 Utilidade Cardinal	126
2.3 O teorema da representação de von Neumann e Morgenstern (vNM)	129
3. Tomando decisões reais	133
3.1 A teoria de Savage	135
3.2 A teoria de Jeffrey	146
4. Implicações mais amplas da teoria da Utilidade Esperada (UE)	153
4.1 Sobre a crença racional	153
4.2 Sobre o desejo racional	155
5. Desafios à teoria da UE	160
5.1 Sobre risco e atitudes de arrependimento	161
5.2 Sobre a completude: crenças e desejos vagos	165
5.3 Desconhecimento	167
6. Decisões sequenciais	169
6.1 Ulisses era racional?	170
6.2 Os axiomas da UE revisitados	174
7. Observações finais	176
Referência bibliográfica	177

(III) Teoria dos Jogos e Ética **185**

1. História	186
2. Funcionalismo	187
3. Problemas com o funcionalismo	189
4. Teoria da barganha e contratarianismo	190
5. Moralidade por acordo	194

6. Alguns problemas com a abordagem contratarianista	196
7. Teoria dos jogos evolucionários e ética	198
8. Algumas observações sobre a abordagem evolucionária	200
9. Algumas aplicações abstratas da teoria dos jogos na ética	201
10. Conclusão	203
Referência bibliográfica	203

(IV) Teoria da Escolha Social	207
1. História da teoria da escolha social	208
1.1 Condorcet	208
1.2 Arrow e sua influência	209
1.3 Borda, Carroll, Black e outros	211
2. Três argumentos formais para a regra da maioria	212
2.1 O conceito de uma regra de agregação	212
2.2 Um argumento procedimental para a regra da maioria	215
2.3 Um argumento epistêmico para a regra da maioria	216
2.4 Um argumento utilitarista para a regra da maioria	219
3. Agregação de preferências	220
3.1 Teorema de Arrow	221
3.2 Regras de agregação de preferências não ditatoriais	223
3.2.1 Relaxando o domínio universal	223
3.2.2 Relaxando a ordem	225
3.2.3 Relaxando o princípio fraco de Pareto	225
3.2.4 Relaxando a independência de alternativas irrelevantes	227
3.3 Teorema de Gibbard-Satterthwaite	230
4. Agregação de bem-estar	233
4.1 Mensurabilidade e comparabilidade interpessoal de bem-estar	234
4.2 A possibilidade de agregação de bem-estar	238
4.3 Outros tópicos	242
5. Agregação de julgamentos	244
5.1 O “paradoxo doutrinário” e o “dilema discursivo”	244
5.2 As condições para julgamentos majoritários inconsistentes	246
5.3 Um modelo geral e um resultado de impossibilidade simples	247
5.4 Teoremas de impossibilidade e possibilidade mais gerais	249
5.5 Regras de agregação de julgamentos não-ditatoriais	252
5.5.1 Relaxando o domínio universal	252

5.5.2 Relaxando a racionalidade coletiva	253
5.5.3 Relaxando sistematicidade / independência	254
6. Outros tópicos	256
Referência bibliográfica	257
(V) Ética dos Negócios	266
1. Variedades da ética dos negócios	267
2. Agência Moral Corporativa	269
3. Os fins e os meios da governança corporativa	271
3.1 Os fins: Primazia dos acionistas ou equilíbrio entre as partes interessadas?	271
3.2 Meios: controle pelos acionistas ou também por outros?	274
4. Estruturas usuais para Ética dos Negócios	276
5. Empresas e consumidores	277
5.1 O que deve ser vendido?	278
5.2 Segurança e confiabilidade do produto	278
5.3 Publicidade	279
5.4 Vendas e Negociações	282
5.5 Atribuição de preços	284
6. Empresas e trabalhadores	286
6.1 Contratação e demissão	286
6.2 Pagamento	288
6.3 Trabalho Significativo	290
6.4 Informantes	291
7. A empresa na sociedade	293
7.1 Responsabilidade social corporativa	294
7.2 Empresas, governos e RSC política	296
7.3 Negócios Internacionais	299
8. O status da ética dos negócios	301
Referência bibliográfica	302
Sobre os tradutores e revisores	316

SOBRE A SÉRIE INVESTIGAÇÃO FILOSÓFICA

A *Série Investigação Filosófica* é uma coleção de livros de traduções de verbetes da *Enciclopédia de Filosofia de Stanford* (*Stanford Encyclopedia of Philosophy*), que se intenciona a servir tanto como material didático, para os professores das diferentes subáreas e níveis da Filosofia, quanto como material de estudo, para a pesquisa e para concursos da área. Nós, professores, sabemos o quão difícil é encontrar bons materiais em português para indicarmos aos estudantes, e há uma certa deficiência na graduação brasileira de Filosofia, principalmente em localizações menos favorecidas, em relação ao conhecimento de outras línguas, como o inglês e o francês. Sendo assim, tentamos suprir essa deficiência, introduzindo essas traduções ao público de Língua Portuguesa, sem nenhuma finalidade comercial, meramente pela glória da Filosofia. Aproveitamos para agradecer a John Templeton Foundation por financiar a publicação de vários dos livros de nossa série, incluindo este, e eximi-la de quaisquer opiniões aqui contidas, as quais são de responsabilidade de seus devidos autores. [*This publication was made possible through a support of a grant from John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.*]

Essas traduções foram todas realizadas por filósofos ou por estudantes de filosofia supervisionados, além de, posteriormente, terem sido revisadas por especialistas nas respectivas áreas. Todas as traduções dos verbetes foram autorizadas pelo querido Prof. Dr. Edward Zalta, editor da *Enciclopédia de Filosofia de Stanford*, razão pela qual o agradecemos imensamente. Sua disposição em contribuir para a ciência brinda os países de Língua Portuguesa com um material filosófico de excelência, disponibilizado gratuitamente no site da Editora da Universidade Federal de Pelotas (UFPel), assim, contribuindo para nosso maior princípio, a ideia de transmissão de conhecimento livre, além de, também, corroborar nossa intenção, a de promover o desenvolvimento da Filosofia em Língua Portuguesa e seu ensino no país. Aproveitamos o ensejo para agradecer, também, ao editor da

UFPel, na figura do Prof. Dr. Juliano do Carmo, que apoiou nosso projeto desde o início. Agradecemos, ainda, a todos os organizadores, tradutores e revisores, que participam de nosso projeto. Sem a dedicação voluntária desses colaboradores, nosso trabalho não teria sido possível. Esperamos, com o início desta Série, abrir as portas para o crescimento desse projeto de tradução e trabalharmos em conjunto pelo crescimento da Filosofia em Língua Portuguesa.

Prof. Dr. Rodrigo Reis Lastra Cid (IF/UNIFAP)
Prof. Dr. Juliano Santos do Carmo (NEFIL/UFPel)
Editores da *Série Investigação Filosófica*

INTRODUÇÃO

“As ideias de economistas e filósofos políticos, tanto quando estão certas quanto quando estão erradas, são mais poderosas do que comumente se pensa.

[...] Homens práticos, que se julgam totalmente isentos de qualquer influência intelectual, costumam ser escravos de algum economista falecido.”

John M. Keynes, *The General Theory of Employment, Interest and Money* (trad. livre)

Há um episódio da série animada *Rick and Morty* (2013-) no qual o protagonista manipula o valor da moeda da “Federação Galáctica”, de modo que todos os sistemas que usam dinheiro passam a registrar o valor “0” para cada unidade monetária. Assim, incertos sobre como continuarão a ser pagos, os funcionários de várias organizações param de trabalhar (ninguém mais está sendo pago para dar ou receber ordens), levando a Federação ao colapso. A situação é exagerada e, por isso, cômica (afinal, trata-se de uma animação); mas imagine o que poderia ocorrer no caso de um colapso dos sistemas de pagamentos globais – é quase como imaginar, como o faz Hobbes, a inexistência do poder político. A razão desses exercícios contrafactuais (o “e se”) de imaginação, comuns na Filosofia, é que podem nos dizer algo sobre a realidade. Há exemplos históricos de instabilidade política associada a crises financeiras e hiperinflação – de tal modo que, em alguns casos, as sociedades afetadas substituem as respectivas moedas fiduciárias nacionais por outros meios de troca (como moedas estrangeiras). Assim, nessa aparentemente inocente piada, esconde-se uma longa discussão teórica sobre a “natureza” e o funcionamento do dinheiro e da autoridade, destacando a centralidade das relações econômicas em sociedades complexas – em particular, a maneira como a possibilidade de influenciar tais relações constitui uma forma de controle social que pode ser complementar (ou concorrente) ao poder político.

No entanto, a despeito da importância da atividade econômica para a sociedade, especialmente na atual era de mercados globais, e da crescente influência da ciência econômica sobre as ciências sociais e humanas, há uma lacuna de

material em português sobre a *Filosofia da Economia*, especialmente, de textos introdutórios aos principais tópicos da área. Sendo assim, preenchê-la é a primeira motivação para a realização deste volume.

Uma das possíveis explicações para essa lacuna é que a Economia inclui temas complexos e recheados de desacordos. De fato, como fica claro já no verbete *Filosofia da Economia*, de Daniel Hausman, economistas investigam fenômenos bastante distintos (desde a análise da tomada de decisão, passando pelo estudo do comportamento de agentes em mercados, até a previsão de variáveis macroeconômicas), com finalidades diferentes (como realizar estimativas, previsões, ou prescrições) e metodologias contrastantes – por exemplo, a recente controvérsia sobre a disseminação do uso de estudos randomizados controlados em economia do desenvolvimento, em detrimento de outros métodos (como a tradicional análise estatística de séries históricas).

Outrossim, parte do problema é que a economia não parece ser uma das questões “clássicas” da história da filosofia. Embora grandes filósofos, de Platão até Hume, tenham abordado pontos atinentes a essa área, isso raramente foi o foco ou o destaque em seus trabalhos – ficando suas ideias para trás à medida que essa ciência progrediu. Ao menos desde Adam Smith (ele mesmo filósofo e economista), quando a economia passou a ser vista como um campo especial e independente das “ciências morais”, invocar pensadores e pensadoras da história da filosofia em discussões econômicas começou a cair em desuso: hoje ninguém consideraria particularmente relevante para a economia monetária, p. ex., um discurso de John Locke sobre como a libra esterlina seria, **por definição**, uma quantidade de prata. Essa demarcação da Economia, e seu distanciamento em relação a outros campos de estudo, como a Filosofia, foi importante para o desenvolvimento de seus métodos científicos próprios. Como os artigos a seguir indicam, entretanto, nunca houve uma separação rígida entre as áreas.

Outra dificuldade é que economistas são, por assim dizer, um tipo bastante especial de cientistas: o que dizem e fazem tem influência direta sobre a produção e a distribuição de recursos na sociedade, mesmo em temas especializados que parecem distantes de nosso cotidiano. Por isso, é improvável que algum dia economistas venham a ser vistos apenas como profissionais meramente competentes, como dentistas – como queria Keynes. Por exemplo, uma política monetária com taxas de juros altas demais tende a favorecer poupadores ricos e a impactar negativamente a produção econômica – podendo assim aumentar o desemprego, a pobreza e a desigualdade; por outro lado, uma política monetária “frouxa” tende

a causar inflação – a qual, além de gerar instabilidade sobre o valor de bens e serviços e as decisões de consumo e investimento ao longo do tempo (como, no limite, mostra exemplo exagerado de *Rick and Morty*), também tende a gerar uma transferência de recursos da sociedade para o governo (i.e., uma taxaço indireta). Dessa forma, um debate eminentemente técnico necessariamente é moldado por decisões e discussões políticas, que são amparadas e discutidas por diversas disciplinas.

É por essa razão, também, que um volume como este precisa ter como seu pilar a interdisciplinaridade. Formamos, assim, um time de tradução e de revisão com profissionais de diversas áreas, como Filosofia do Direito, Filosofia Política, Filosofia da Ciência, Ética Aplicada e Economia. Os temas explorados nesses verbetes guardam relação com diversas áreas da Filosofia – o que é especialmente evidente no caso da Filosofia Política, a qual há muito se tem beneficiado de uma frutífera interação com a Economia (como qualquer leitor de John S. Mill ou de John Rawls já deve ter notado). Mas tal influência também é pervasiva em outras áreas, particularmente na filosofia contemporânea; por exemplo, a Teoria da Decisão tem sido cada vez mais estudada por pesquisadores ligados à Teoria da Ação, à Filosofia da Mente e à Filosofia da Ciência (em particular por meio de sua influência sobre a Filosofia da Probabilidade e da Estatística) – como pretendia o precursor da área, o brilhante filósofo-economista-matemático Frank Ramsey.

Os verbetes deste primeiro volume, conforme já mencionado, foram selecionados de forma a abordar questões filosóficas mais abstratas presentes nos próprios fundamentos da Ciência Econômica; além do panorama geral fornecido por Hausman no verbete *Filosofia da Economia*, incluímos: *Teoria da Decisão*, *Teoria dos Jogos e Ética*, *Teoria da Escolha Social*, e *Ética dos Negócios*. Embora possa parecer que tratem de questões relativamente específicas (e o formalismo presente em alguns dos verbetes possa assustar iniciantes), na realidade, estes temas são básicos para a análise das decisões de agentes individuais (sozinhos, ou em interações estratégicas) ou coletivos (como sociedades e instituições). Esperamos que, vocês, nossos leitores e leitoras, estudantes ou especialistas em economia, filosofia e teoria política, dentre outras áreas, encontrem aqui temas que os levem a refletir.

Agradecimentos são a moeda com que se quitam dívidas morais. Dessa forma, somos gratos, primeiramente, à própria *Stanford Encyclopedia of Philosophy* e a seu editor, Edward N. Zalta, pela existência deste valioso recurso, disponível online para todos, e por autorizarem a tradução de verbetes selecionados para a Língua Portuguesa. Aos respectivos autores, pelo brilhante trabalho de pesquisa, análise, escrita e sumarização. À organização do projeto de tradução, ao *Grupo de*

Pesquisa Investigação Filosófica, e ao Prof. Dr. Rodrigo Reis Lastra Cid – pela liderança no projeto da Série Investigação Filosófica. À *John Templeton Foundation*, cujo apoio financeiro, além de prestigiar o projeto, constituiu um importante incentivo a este trabalho. Aos colegas do *Grupo de Estudos em Filosofia da Economia* – um projeto paralelo desenvolvido, ao longo dessas traduções, para discutir estes temas, encabeçado pelo Prof. Dr. André Pontes (um dos organizadores destes volumes). E, principalmente, aos tradutores e revisores que aceitaram dar vida ao projeto.

Também somos gratos a vocês, leitores e leitoras. Desejamos que possam encontrar nessas páginas ao menos um pouco da excitação e do estímulo intelectual que nós mesmos tivemos ao ler estes verbetes ou ao ter o primeiro contato com a área, a de desbravar um novo campo, ou de colocar ordem a ideias e conceitos e de descobrir relações surpreendentes entre eles, ou de simplesmente aprender sobre uma parte importante do mundo onde vivemos. Em especial, gostaríamos de contribuir com sua formação intelectual, e que isso seja benéfico não só para você mesmo, mas também para os demais. Afinal, talvez a principal lição da filosofia da economia seja perceber o quanto podemos contribuir uns para com os outros, gerando relações de interdependência e de responsabilidade mútua, das quais raramente nos damos conta: perceber, assim, que os sistemas de divisão de trabalho baseados em especialização e de trocas auto interessadas de produtos e serviços entre estranhos num mercado, mediadas pelo “vil metal”, na realidade, só são possíveis, enquanto práticas sociais continuadas, na medida em que se inserem numa rede de relações de cooperação. Dessa forma, mesmo o mais altruísta dos consumidores, nas últimas décadas, emitiu gases de efeito estufa que hoje contribuem para a seca e a fome que assolam o sul de Madagascar; e mesmo o mais egoísta dos especuladores financeiros de hoje deve seu emprego e riqueza ao fato de que há outras pessoas dispostas a poupar (algumas das quais não estão mais aqui), a trabalhar e a consumir (algumas das quais ainda não existem).

Essa lição é a razão definitiva deste volume.

Organizadores

André Nascimento Pontes,
Mariana Kuhn de Oliveira e
Ramiro de Ávila Peres

(I) Filosofia da Economia*

Daniel M. Hausman (Universidade de Wisconsin-Madison)

Tradução: Ramiro de Ávila Peres (BCB)

Revisão: Mariana Kuhn de Oliveira (CEBRAP)

A “Filosofia da Economia” consiste em indagações sobre (a) a escolha racional, (b) a avaliação dos resultados, instituições e processos econômicos, e (c) a ontologia dos fenômenos econômicos e as possibilidades de adquirir conhecimento deles. Embora essas indagações se sobreponham de muitas maneiras, é útil dividir a filosofia da economia desta forma em três temas que podem ser considerados, respectivamente, como ramos da teoria da ação, da ética (ou filosofia social e política normativa) e da filosofia da ciência. As teorias econômicas da racionalidade, do bem-estar e da escolha social defendem teses de filosofia substantiva frequentemente informadas pela literatura filosófica relevante e de evidente interesse para aqueles interessados em teoria da ação, psicologia filosófica e filosofia social e política. A economia é de particular interesse para os interessados em epistemologia e filosofia

* HAUSMAN, D. M. Economics. In: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Fall Edition. Stanford, CA: The Metaphysics Research Lab, 2018. Disponível em: <https://plato.stanford.edu/archives/fall2018/entries/economics/>.

The following is the translation of the entry on “Philosophy of Economics” by Daniel M. Hausman, in the Stanford Encyclopedia of Philosophy. The translation follows the version of the entry in the SEP’s archives at <<https://plato.stanford.edu/archives/fall2018/entries/economics/>>. This translated version may differ from the current version of the entry, which may have been updated since the time of this translation. The current version is located at a <<https://plato.stanford.edu/entries/economics/>>. We’d like to thank the Editors of the Stanford Encyclopedia of Philosophy, mainly Prof. Dr. Edward Zalta, for granting permission to translate and to publish this entry.

da ciência tanto por suas peculiaridades detalhadas quanto por possuir muitas das características evidentes das ciências naturais, enquanto seu objeto consiste em fenômenos sociais.

1. Introdução: O que é Economia?

Tanto a definição quanto o domínio preciso da economia são temas de controvérsia para a filosofia da economia. À primeira vista, as dificuldades na definição de economia podem não parecer sérias. A economia está, afinal, preocupada com aspectos da produção, troca, distribuição e consumo de produtos e serviços. Mas essa afirmação e os termos que ela contém são vagos; e pode-se argumentar que a economia é relevante para muito mais. Ajuda abordar historicamente a questão: “O que é economia?”, antes de passarmos aos comentários sobre as características contemporâneas da disciplina.

1.1 O surgimento da ciência econômica e das economias

A reflexão filosófica sobre economia é antiga, mas a concepção da economia como objeto de estudo distinto remonta apenas ao século XVIII. Aristóteles aborda alguns problemas que a maioria reconheceria como pertencentes à economia, principalmente problemas de administração doméstica. Os filósofos escolásticos abordaram questões éticas relativas ao comportamento econômico e condenaram a usura, ou seja, a tomada de juros pelo dinheiro. Com a crescente importância do comércio e dos estados-nação no início do período moderno, filósofos e panfletários “mercantilistas” preocupavam-se em grande parte com a balança comercial e a regulamentação da moeda. Havia um reconhecimento crescente das complexidades da gestão financeira do Estado e da possibilidade de que a forma como o Estado tributava e agia influenciava a produção de riqueza.

No início do período moderno, aqueles que refletiram sobre as fontes de riqueza de um país reconheceram que a colheita anual, as quantidades de bens fabricados e os produtos das minas e da pesca dependem de fatos sobre a natureza, a mão-de-obra individual e a empresa, ferramentas e o que chamaríamos de “bens de capital”, e regulamentações estatais e sociais. O comércio também parecia vantajoso, pelo menos se os termos fossem suficientemente bons. Não foi preciso

um salto conceitual para reconhecer que a produção e a agricultura poderiam ser melhoradas e que alguns impostos e tarifas poderiam ser menos prejudiciais às atividades produtivas do que outros. Mas formular a ideia de que existe uma coisa como a “economia” com regularidades que podem ser investigadas requer um passo ousado. Para que haja um objeto de investigação, é preciso que haja regularidade na produção e na troca; e para que a investigação seja não-trivial, essas regularidades devem ir além do que é óbvio para os produtores, consumidores e os próprios trocadores. Somente no século XVIII, mais claramente ilustrado pelo trabalho de Cantillon, dos Hume, e especialmente de Adam Smith (*vide verbete sobre a filosofia moral e política de Smith no site da Stanford Encyclopedia of Philosophy*), se de interações que produzem e distribuem bens de consumo e os recursos e ferramentas que os produzem (BACKHOUSE, 2002).

Crucial para a possibilidade de um objeto social de investigação científica é a ideia de traçar as consequências não intencionais das ações intencionais dos indivíduos. Assim, por exemplo, Hume traça o aumento dos preços e o aumento temporário da atividade econômica que se segue a um aumento da moeda para as percepções e ações dos indivíduos que primeiro gastam a moeda adicional (1752). Ao gastar seu ouro adicional importado do exterior, comerciantes não pretendem aumentar o nível de preços. No entanto, é isso que eles fazem. Adam Smith expande e aperfeiçoa essa percepção e oferece uma sistemática *Investigação sobre a Natureza e as Causas da Riqueza das Nações*. Do seu relato do fim do feudalismo (1776, Livro II, cap. 4) a sua famosa discussão sobre a mão invisível, Smith enfatiza as consequências involuntárias. “[E]le pretende apenas o próprio ganho; e ele é, como em muitos outros casos, conduzido por uma mão invisível para promover um fim que não fazia parte de sua intenção. Nem sempre é pior para a sociedade que isso não fizesse parte dela. Ao perseguir seu próprio interesse, o indivíduo frequentemente promove o da sociedade mais efetivamente do que quando realmente pretende promovê-lo” (1776, Livro IV, cap. 2). A existência de regularidades, que são consequências não intencionais das escolhas individuais, dá origem a um objeto de investigação científica.

Pode-se distinguir o domínio da economia do domínio de outras investigações científicas sociais, seja especificando algum conjunto de fatores causais, seja especificando alguma gama de fenômenos. Os fenômenos que preocupam os economistas são a produção, o consumo, a distribuição e a troca – especialmente através dos mercados. Mas como tantos fatores causais diferentes são relevantes para estes, incluindo as leis da termodinâmica, metalurgia, geografia e normas

sociais, mesmo as leis que regem a digestão, a economia não pode ser distinguida de outras indagações **apenas** pelos fenômenos que estuda. É necessária alguma referência a um conjunto de fatores causais centrais. Assim, por exemplo, sustenta que, “A economia política [preocupa-se] com os fenômenos do estado social que ocorrem em consequência da busca da riqueza”. Ele abstrai completamente qualquer outra paixão ou motivo humano, exceto aqueles que podem ser considerados como princípios perpetuamente antagônicos ao desejo de riqueza, nomeadamente a aversão ao trabalho, e o desejo do presente gozo de indulgências dispendiosas” (1843, Livro VI, cap. 9, seq. 3). Na visão de Mill, a economia preocupa-se principalmente com as consequências da busca individual de riqueza tangível, embora leve em conta motivos menos significativos, como a aversão ao trabalho.

Mill toma como certo que os indivíduos agem racionalmente em sua busca de riqueza e luxo e em evitar o trabalho, e não de forma desarticulada ou errática, mas ele não possui uma teoria de consumo, ou teoria explícita de escolha econômica racional, e sua teoria de alocação de recursos é bastante tênue. Essas lacunas foram gradualmente preenchidas durante a chamada revolução neoclássica ou marginalista, que vinculou a escolha de algum objeto de consumo (e seu preço) não à sua utilidade total, mas à sua utilidade marginal. Por exemplo, a água é obviamente extremamente útil, mas em boa parte do mundo é abundante o suficiente para que outro copo importe mais ou menos para um agente. Portanto, a água é barata. Os primeiros economistas “neoclássicos”, como William Stanley Jevons, sustentavam que os agentes fazem escolhas de consumo de forma a maximizar sua própria felicidade (1871). Isto implica que eles distribuem seus gastos para que o valor de um dólar de água ou mingau ou estofado faça a mesma contribuição para sua felicidade. A “utilidade marginal” do valor de um dólar de cada bem é a mesma.

No século XX, os economistas despojaram esta teoria de sua roupagem hedonista (PARETO, 1909; HICKS; ALLEN, 1934). Em vez de supor que todas as escolhas de consumo podem ser classificadas pelo quanto promovem a felicidade de um agente, os economistas se concentraram na própria ordenação. Tudo o que eles supõem em relação às avaliações é que os agentes são capazes de classificar consistentemente as alternativas que eles enfrentam. Isto equivale a supor primeiro que as classificações são completas – ou seja, para quaisquer duas alternativas **x** e **y** que o agente considere, ou o agente coloca **x** acima de **y** (prefere **x** a **y**), ou o agente prefere **y** a **x**, ou o agente é indiferente. Em segundo lugar, os economistas supõem que as classificações de alternativas (preferências) do agente são transitivas. Dizer que as preferências de um agente são transitivas é afirmar que se o agente

prefere x a y e y a z , então o agente prefere x a z , com afirmações semelhantes relativas à indiferença e combinações de indiferença e preferência. Embora existam outras condições técnicas para estender a teoria a infinitos conjuntos de alternativas e para captar mais condições plausíveis de racionalidade concernente a apostas, economistas geralmente subscrevem uma visão dos agentes racionais como possuindo pelo menos preferências completas e transitivas e como escolhendo entre as alternativas viáveis, as que mais preferem. Na teoria da preferência revelada, economistas têm tentado, sem sucesso, eliminar toda referência à preferência subjetiva ou definir preferência em termos de escolhas (SAMUELSON, 1947; HOUTTHAKER, 1950; LITTLE, 1957; SEN, 1971, 1973; HAUSMAN, 2012, cap. 3).

Ao esclarecer a visão de racionalidade que caracteriza os agentes econômicos, os economistas têm, em sua maioria, continuado a distinguir a economia das demais investigações sociais pelo conteúdo dos motivos ou preferências com os quais ela se preocupa. Assim, embora as pessoas possam buscar a felicidade por meio da abstinência, ou possam preferir racionalmente sacrificar todos os seus bens mundanos a uma causa política, os economistas supõem que tais preferências são raras e sem importância para a economia. Os economistas estão preocupados com os fenômenos decorrentes da racionalidade associada a um desejo de riqueza e de quantidades maiores de bens e serviços.

Os economistas têm flertado com uma caracterização menos substantiva da motivação individual e com uma visão mais expansiva do domínio da economia. Em sua influente monografia, *Um Ensaio sobre a Natureza e a Significância da Ciência Econômica*, Lionel Robbins (1932, p. 15) definiu a economia como “a ciência que estuda o comportamento humano como uma relação entre fins e meios escassos que têm usos alternativos”. Segundo Robbins, a economia não se preocupa com a produção, troca, distribuição ou consumo como tais. Ela se preocupa antes com um aspecto de toda ação humana. A definição de Robbins ajuda a entender os esforços para aplicar conceitos, modelos e técnicas econômicas a outros assuntos, como a análise do comportamento de voto e legislação, mesmo quando a economia mantém sua conexão com um domínio tradicional.

1.2 A economia contemporânea e suas diversas escolas

A economia contemporânea é diversa. Há muitas escolas e muitos ramos. Mesmo a chamada economia “ortodoxa” ou “*mainstream*” tem muitas variantes.

Uma fração da economia *mainstream* é altamente teórica, sua maior parte é, contudo, aplicada e se baseia em teoria rudimentar. O trabalho teórico e aplicado pode ser distinguido tanto como microeconomia quanto como macroeconomia. Há também um terceiro ramo, a econometria, que é dedicado à estimativa empírica, elaboração e, em certa medida, teste de modelos microeconômicos e macroeconômicos (*vide* SUMMERS, 1991; HOOVER, 1994).

A microeconomia foca nas relações entre os indivíduos (sendo que empresas e domicílios frequentemente contam como indivíduos honorários, e pouco se fala sobre as idiossincrasias da demanda de indivíduos particulares). Os indivíduos têm preferências completas e transitivas que regem suas escolhas. Os consumidores preferem mais produtos a menos e têm “taxas marginais decrescentes de substituição” – ou seja, pagarão menos por unidades de um produto quando já tiverem muitas delas do que quando tiverem poucas. As empresas tentam maximizar os lucros diante da diminuição dos retornos: mantendo fixos todos os insumos na produção, exceto um, a produção aumenta quando há mais do insumo restante, mas a uma taxa decrescente. Os economistas idealizam e supõem que em mercados competitivos, empresas e indivíduos não podem influenciar os preços, mas os economistas também estão interessados em interações estratégicas, nas quais as escolhas racionais de indivíduos separados são interdependentes. A teoria dos jogos, que se dedica ao estudo das interações estratégicas, é de importância crescente na economia. Os economistas modelam o resultado das atividades maximizadoras de lucro das empresas e as tentativas dos consumidores de satisfazer suas preferências de forma otimizada como um **equilíbrio** em que não há excesso de demanda em nenhum mercado. O que isto significa é que qualquer pessoa que queira comprar qualquer coisa a preço de mercado é capaz de fazer isso. Não há excesso de demanda, e, a menos que um bem seja gratuito, não há excesso de oferta.

A macroeconomia lida com as relações entre os agregados econômicos, tais como as relações entre a oferta de moeda e a taxa de juros ou a taxa de crescimento, com foco especial nos problemas relativos ao ciclo econômico e à influência da política monetária e fiscal sobre os resultados econômicos. Muitos dos principais economistas gostariam de unificar a macroeconomia e a microeconomia, mas poucos economistas estão satisfeitos com as tentativas feitas nesse sentido, especialmente por meio dos chamados “agentes representativos” (KIRMAN 1992; HOOVER, 2001a). A macroeconomia é imediatamente relevante para a política econômica e, portanto (sem surpresas), sujeita a controvérsias muito mais acaloradas (e politicamente carregadas) do que a microeconomia ou a econometria. As escolas

de macroeconomia incluem Keynesianos (e neo-Keyesianos), monetaristas, “nova economia clássica”. A esse respeito, confira sobre a teoria das expectativas racionais em Begg (1982), Carter e Maddock (1984), Hoover (1988), Minford e Peel (1983) e as teorias do “ciclo real”, em Kydland e Prescott (1991, 1994) e Sent (1998).

Os ramos da economia *mainstream* também se dedicam a questões específicas relativas ao crescimento, finanças, emprego, agricultura, habitação, recursos naturais, comércio internacional, e assim por diante. Na economia ortodoxa, há também muitas abordagens diferentes, como a **teoria da agência** (JENSEN; MECKLING, 1976; FAMA, 1980), a **escola de Chicago** (BECKER, 1976), ou a **teoria da escolha pública** (BRENNAN; BUCHANAN, 1985; BUCHANAN, 1975). Estas tratam de questões relativas a incentivos dentro das empresas e famílias e as formas pelas quais as instituições orientam as escolhas.

Embora a economia *mainstream* seja dominante e exija a maior atenção, existem muitas outras escolas. Os **economistas austríacos** aceitam visões ortodoxas de escolhas e restrições, mas enfatizam a incerteza e questionam se devemos considerar os resultados como equilíbrios e são céticos quanto ao valor da modelagem matemática (BUCHANAN; VANBERG, 1989; DOLAN, 1976; KIRZNER, 1976; MISES, 1949, 1978, 1981; ROTHBARD, 1957; WISEMAN, 1983; BOETTKE, 2010; HOLCOMBE, 2014; NELL, 2014a, 2014b, 2017; BOETTKE; COYNE, 2015; HAGEDORN, 2015; HORWITZ, 2015; DEKKER, 2016; LINSBICHLER, 2017).

Os **economistas institucionalistas** tradicionais questionam o valor da teorização geral abstrata e enfatizam conceitos evolutivos (DUGGER, 1979; WILBER; HARRISON, 1978; WISMAN; ROZANSKY, 1991; HODGSON, 2000, 2013, 2016; HODGSON; KNUDSEN, 2010; DELORME, 2010; RICHTER, 2015). Eles enfatizam a importância de generalizações sobre normas e comportamentos dentro de instituições particulares. A pesquisa aplicada em economia institucional é por vezes muito semelhante à economia ortodoxa aplicada. Trabalhos mais recentes em economia, também chamado de institucionalista, buscam explicar características das instituições, enfatizando os custos das transações, a inevitável incompletude dos contratos e os problemas que os *principais* enfrentam no monitoramento e direção de seus agentes (COASE, 1937; WILLIAMSON, 1985; MÄKI *et al.*, 1993; NORTH, 1990; BROUSSEAU; GLACHANT, 2008).

Economistas marxianos e socialistas tradicionalmente articularam e desenvolveram as teorias econômicas de Karl Marx, mas, recentemente, muitos economistas socialistas revisaram conceitos e temas marxianos tradicionais com ferramentas emprestadas da teoria econômica ortodoxa (MORISHIMA, 1973;

ROEMER, 1981, 1982; BOWLES, 2012; PIKETTY, 2014; LEBOWITZ, 2015; AUERBACH, 2016; BECKERT, 2016; JACOBS; MAZZUCATO, 2016).

Há também os **socioeconomistas**, que se preocupam com as normas que regem as escolhas (ETZIONI, 1988, 2018), os **economistas comportamentais**, que estudam as nuances de comportamentos de escolha (WINTER, 1962; THALER, 1994; BEN NER; PUTTERMAN, 1998; KAHNEMAN; TVERSKY, 2000; CAMERERER, 2003; CAMERERER; LOEWENSTEIN, 2003; CAMERERER *et al.*, 2003; LOEWENSTEIN, 2008; THALER; SUNSTEIN, 2008; SAINT-PAUL, 2011; OLIVER, 2013), **pós-keynesianos**, que se inspiram no trabalho de Keynes e, especialmente, em sua ênfase na demanda (DOW, 1985; KREGEL, 1976; HARCOURT; KRIESLER, 2013; ROCHON; ROSSI, 2017), **economistas evolucionistas**, que enfatizam a importância das instituições (WITT, 2008; HODGSON; KNUDSEN, 2010; VROMEN, 2009; HODGSON, 2013, 2016; CARSTEN, 2013; DOPFER; POTTS, 2014; WILSON; KIRMAN, 2016), **neo-Ricardianos**, que enfatizam as relações entre classes econômicas (SRAFFA, 1960; PASINETTI, 1981; RONCAGLIA, 1978), e até mesmo **neuroeconomistas**, que estudam aspectos neurológicos de comportamentos de escolha (CAMERERER, 2007; CAMERERER *et al.*, 2005; CAMERERER *et al.*, 2008; GLIMCHER *et al.*, 2008; LOEWENSTEIN *et al.*, 2008; RUSTICINNI, 2005, 2008; GLIMCHER, 2010). A economia não é um empreendimento homogêneo.

2. Seis problemas metodológicos centrais

Embora os diferentes ramos e escolas de economia levantem uma grande variedade de questões epistemológicas e ontológicas relativas à economia, seis problemas têm sido centrais para a reflexão metodológica (neste sentido filosófico) relativa à economia.

2.1 Economia positiva versus normativa

Os formuladores de políticas públicas buscam a economia para orientar seu trabalho, e parece inevitável que até mesmo as questões mais esotéricas da economia teórica possam ter influência sobre os interesses materiais de algumas pessoas. A medida em que a economia se sustenta e pode ser influenciada por preocupações normativas, levanta questões metodológicas sobre as relações entre

uma ciência **positiva**, relacionada aos “fatos”, e uma investigação **normativa**, referente a valores e sobre como deve ser. A maioria dos economistas e metodologistas acredita que há uma distinção razoavelmente clara entre fatos e valores, entre o que é e o que deve ser, e acredita que a maior parte da economia deve ser considerada como uma ciência positiva que ajuda os formuladores de políticas públicas a escolher os meios para atingir seus fins, embora não suporte a escolha dos próprios fins.

Esta visão é questionável por várias razões (MONGIN, 2006; HAUSMAN, MCPHERSON; SATZ, 2017). Primeiro, economistas têm que interpretar e articular as especificações incompletas de objetivos e restrições fornecidas pelos formuladores de políticas (MACHLUP, 1969b). Segundo, a “ciência” econômica é uma atividade humana, e como todas as atividades humanas, ela é governada por valores. Esses valores não precisam ser iguais aos valores que influenciam a política econômica, mas é discutível se os valores que governam a atividade dos economistas podem ser claramente distinguidos dos valores que governam os formuladores de políticas públicas. Terceiro, grande parte da economia é construída em torno de uma teoria normativa da racionalidade. Pode-se questionar se os valores implícitos em tais teorias são claramente distinguíveis dos valores que governam as políticas. Por exemplo, pode ser difícil ter uma visão maximizadora da racionalidade individual, ao mesmo tempo em que se insiste que a política social deve resistir a maximizar o crescimento, a riqueza ou o bem-estar em nome da liberdade, dos direitos ou da igualdade. Quarto, a visão das pessoas sobre o que é certo e errado é, de fato, influenciada por suas crenças sobre como as pessoas de fato se comportam. Há evidências de que o estudo de teorias que retratam os indivíduos como autointeressados leva as pessoas a considerar o comportamento autointeressado de forma mais favorável e a se tornarem mais autointeressadas (MARWELL; AMES, 1981; FRANK *et al.*, 1993). Finalmente, os julgamentos das pessoas são obscurecidos por seus interesses. Uma vez que as teorias econômicas se apoiam de forma tão central nos interesses das pessoas, é provável que haja preconceitos ideológicos no trabalho da disciplina (MARX, 1867, Prefácio). Positivistas e normativistas estão especialmente interligados no campo da economia, pois os economistas não são todos pesquisadores e professores. Além disso, os economistas trabalham como comentaristas e, por assim dizer, como “mercenários”, cujos salários dependem de eles chegarem às conclusões que seus empregadores desejam. A amarga polêmica sobre as respostas de política macroeconômica à grande recessão iniciada em 2008 atesta a influência da ideologia.

2.2 Razões versus causas

A microeconomia teórica ortodoxa é tanto uma teoria de escolhas racionais quanto uma teoria que explica e prevê resultados econômicos. Como praticamente todas as teorias econômicas que discutem escolhas individuais tomam os indivíduos como agindo por razões, e, portanto, de alguma forma racional, questões sobre o papel que as visões da racionalidade e que as razões devem desempenhar na economia são de importância geral. Os economistas normalmente se preocupam com os resultados agregados das escolhas individuais e não com as ações de determinados indivíduos, mas suas teorias de fato oferecem tanto explicações causais para o porquê de os indivíduos escolherem como o fazem quanto relatos das **razões** de suas escolhas. Veja também os verbetes sobre e , no site da SEP, em inglês.

As explicações em termos de razões têm várias características que as distinguem das explicações em termos de causas. Razões pretendem **justificar** as ações que elas explicam, e de fato as “razões externas” (WILLIAMS, 1981) apenas justificam a ação, sem pretender explicá-la. As razões podem ser avaliadas, e são sensíveis às críticas. As razões, ao contrário das causas, devem ser inteligíveis para aqueles para quem elas são razões. Por fundamentos como estes, muitos filósofos têm questionado se as explicações da ação humana podem ser explicações causais (VON WRIGHT, 1971; WINCH, 1958). No entanto, apenas fornecer uma razão – mesmo uma razão extremamente boa – falha em explicar a ação de um agente, se a razão não foi de fato “efetiva”. Alguém pode, por exemplo, começar a frequentar a igreja regularmente e dar como motivo uma preocupação com a salvação. Mas outros podem suspeitar que esse agente esteja se enganando e que a filha atraente do ministro é, na verdade, responsável pelo seu renovado interesse pela religião. Donald Davidson (1963) argumentou que o que distingue as razões que explicam uma ação das razões que não a explicam é que as primeiras também são causas da ação. Embora o relato da racionalidade dentro da economia difira em alguns aspectos da que as pessoas invocam tacitamente nas explicações cotidianas das ações, muitas das mesmas questões se colocam para ambas (ROSENBERG, 1976, cap. 5, 1980; HAUSMAN, 2012).

Uma diferença adicional entre explicações em termos de razões e explicações em termos de causas, que alguns economistas têm enfatizado, é que as crenças e preferências que explicam ações podem depender de erros e ignorância (KNIGHT, 1935). Como uma primeira aproximação, os economistas podem abstrair-se de tais dificuldades causadas pela da crença e do desejo. Assim, eles muitas vezes assumem

que as pessoas têm informações perfeitas sobre todos os fatos relevantes. Dessa forma, os teóricos não precisam se preocupar com o que são as crenças das pessoas (se as pessoas têm informações perfeitas, então elas acreditam e esperam o que quer que sejam os fatos). Mas, uma vez que se vá além dessa primeira aproximação, surgem dificuldades que não têm paralelo nas ciências naturais. A escolha depende de como as coisas parecem “de dentro”, o que pode ser muito diferente do estado atual das coisas. Considere, por exemplo, o mercado de ações. O valor “verdadeiro” de uma ação depende dos lucros futuros da empresa, que são, naturalmente, incertos. Em 2006 os preços dos imóveis nos Estados Unidos estavam extremamente inflados. Mas se eles estavam “muito altos” dependia, pelo menos no curto prazo, do que as pessoas acreditavam. Eram excelentes investimentos se se pudessem vendê-los a outros que estivessem dispostos a pagar ainda mais por eles. Os economistas discordam sobre o quão significativa é essa subjetividade. Membros da escola austríaca argumentam que essas diferenças são de grande importância e distinguem fortemente a teorização sobre economia da teorização sobre qualquer das ciências naturais (BUCHANAN; VANBERG, 1989; VON MISES, 1981).

2.3 Naturalismo científico social

De todas as ciências sociais, a economia é a que mais se assemelha às ciências naturais. Teorias econômicas têm sido axiomatizadas, e artigos e livros de economia estão cheios de teoremas. De todas as ciências sociais, apenas a economia ostenta um Prêmio Nobel “ersatz”. A economia é, portanto, um caso para testar aqueles interessados na extensão das semelhanças entre as ciências naturais e as ciências sociais. Aqueles que se perguntaram se as ciências sociais devem diferir fundamentalmente das ciências naturais parecem ter se preocupado principalmente com três questões:

- i. Existem diferenças fundamentais entre a estrutura ou conceitos de teorias e nas ciências naturais e sociais? Algumas dessas questões já foram mencionadas na discussão acima de razões versus causas.
- ii. Existem diferenças fundamentais nos objetivos? Filósofos e economistas argumentam que, além ou

ao invés dos objetivos preditivos e explicativos das ciências naturais, as ciências sociais devem ter como objetivo nos proporcionar compreensão. Weber e outros argumentam que as ciências sociais devem nos proporcionar um entendimento “de dentro”, que devemos ser capazes de empatizar com as reações dos agentes e considerar o que acontece como “compreensível” (WEBER, 1904; KNIGHT, 1935; MACHLUP, 1969a). Isto (e o reconhecimento intimamente relacionado que as explicações citam razões e não apenas causas) parece introduzir um elemento de subjetividade nas ciências sociais que não se encontra nas ciências naturais.

iii. Devido à importância das escolhas humanas (ou talvez do livre arbítrio), os fenômenos sociais são demasiado irregulares para serem capturados dentro de um quadro de leis e teorias? Dado o livre arbítrio humano, talvez o comportamento humano seja intrinsecamente imprevisível e não esteja sujeito a quaisquer leis. Mas há, de fato, muitas regularidades na ação humana, e dada a enorme complexidade causal que caracteriza alguns sistemas naturais, as ciências naturais também devem lidar com muitas irregularidades.

2.4 Abstração, idealização e cláusulas *ceteris paribus* em economia

A economia levanta questões quanto à legitimidade da abstração severa e da idealização. Por exemplo, os modelos econômicos *mainstream* muitas vezes estipulam que todos são perfeitamente racionais e têm informações perfeitas ou que as mercadorias são infinitamente divisíveis. Tais afirmações são exageros e são claramente falsas. Outras escolas de economia podem não empregar idealizações que são extremas dessa forma, mas não há como fazer economia se não se está disposto a simplificar drasticamente e abstrair-se de muitas complicações. Quanta simplificação, idealização, abstração ou “isolamento” (MÄKI, 2006) são legítimas?

Além disso, uma vez que os economistas tentam estudar os fenômenos econômicos como constituindo um domínio à parte, influenciados apenas por um

pequeno número de fatores causais, as afirmações da economia são verdadeiras apenas *ceteris paribus* – ou seja, só são verdadeiras se não houver interferências ou causas impeditivas. O que são cláusulas *ceteris paribus* e quando são legítimas na ciência (se é que o são)? As questões relativas às cláusulas *ceteris paribus* estão intimamente relacionadas a questões de simplificação e idealizações, pois uma forma de simplificar é supor que as várias causas ou interferências perturbadoras estão inativas e explorar as consequências de um pequeno número de fatores causais. Essas questões e a questão relacionada de quão bem sustentada é a economia pela evidência têm sido as questões centrais na metodologia econômica. Elas serão discutidas mais adiante, principalmente na .

2.5 Causas em economia e econometria

Muitas generalizações importantes em economia são afirmações causais. Por exemplo, a lei da demanda afirma que um aumento de preço (*ceteris paribus*) diminuirá a quantidade demandada. (Não se trata apenas de afirmar uma relação inversa entre preço e demanda. Quando a demanda aumenta por algum outro motivo, como a mudança de gosto, o preço **aumenta**). Os econométricos também têm se preocupado profundamente com as possibilidades de determinação de relações causais a partir de evidências estatísticas e com a relevância das relações causais para a possibilidade de estimação consistente dos valores dos parâmetros. Como as preocupações com as consequências das políticas alternativas são tão centrais para a economia, a investigação causal é inevitável.

Antes da década de 1930, os economistas estavam geralmente dispostos a usar a linguagem causal de forma explícita e literal, apesar de algumas preocupações de que poderia haver um conflito entre a análise causal das mudanças econômicas e os tratamentos “estáticos comparativos” dos estados de equilíbrio. Alguns economistas também se preocupavam que pensar em termos de causas não fosse compatível com o reconhecimento da multiplicidade e da mutualidade de determinação em equilíbrio econômico. No ambiente intelectual antimetafísico das décadas de 1930 e 1940 (do qual o positivismo lógico era pelo menos sintomático), qualquer menção de causalidade tornou-se suspeita, e os economistas comumente pretendiam evitar conceitos causais. A consequência foi que eles deixaram de refletir cuidadosamente sobre os conceitos causais que eles continuaram implicitamente a invocar (HAUSMAN, 1983, 1990; HELM, 1984; RUNDE, 1998). Por exemplo, em vez de formular a lei

da demanda em termos das consequências causais das mudanças de preço para a quantidade demandada, os economistas tentaram limitar-se a discutir a função matemática relacionada a preço e quantidade demandada. Houve exceções importantes (HAAVELMO, 1944; SIMON, 1953; WOLD, 1954) e, durante a geração passada, este estado de coisas mudou drasticamente.

Por exemplo, em *Causality in Macroeconomics* [Causalidade em Macroeconomia] (2001b) Kevin Hoover desenvolve métodos viáveis para investigar questões causais em larga escala, tais como se mudanças na oferta de moeda (**M**) causam mudanças na taxa de inflação **P** ou acomodam mudanças em **P** que de outra forma são causadas. Se mudanças em **M** causam mudanças em **P**, então a distribuição condicional de **P** sobre **M** deve permanecer estável com mudanças exógenas em **M**, mas deve mudar com mudanças exógenas em **P**. Hoover argumenta que a investigação histórica, apoiada por investigação estatística, pode justificar a conclusão de que algumas mudanças particulares em **M** ou **P** foram exógenas. Pode-se então determinar a direção causal, examinando a estabilidade das distribuições condicionais. Os econometristas têm dado contribuições vitais para o renascimento contemporâneo do interesse filosófico na noção de causalidade. Além da obra de Hoover, confira, por exemplo, Geweke (1982), Granger (1969, 1980), Cartwright (1989), Sims (1977), Zellner e Aigner (1988), Pearl (2000), Spirtes, Glymour e Scheines (2001).

Uma maneira relativamente segura de determinar as relações causais é através de experimentos aleatórios controlados. Se os experimentadores classificam os sujeitos aleatoriamente em grupos experimentais e de controle e variam apenas um fator, então, a menos que por azar os dois grupos difiram de alguma forma desconhecida, as mudanças nos resultados, dadas as características comuns dos grupos de controle e de tratamento, devem ser devidas à diferença de um fator. De fato, no caso das variáveis quantitativas, pode-se calcular os efeitos causais médios (DEATON, 2010). Isto torna os estudos controlados aleatórios muito atraentes, embora nenhuma panaceia, uma vez que os grupos de tratamento e de controle podem não ser representativos da população na qual os formuladores de políticas públicas esperam aplicar as conclusões causais, e as consequências causais da intervenção podem diferir entre diferentes subgrupos dentro dos grupos de controle e de tratamento (WORRALL, 2007; CARTWRIGHT; HARDIE, 2013).

Por razões tanto práticas quanto éticas, muitas vezes é difícil realizar experimentos em economia (embora, como será discutido na seção 4.5, isso está longe de ser impossível). Contudo, com alguma engenhosidade e com muito

entusiasmo pela experimentação, o que não era o caso antes, os economistas estão realizando experimentos com muito mais frequência tanto no laboratório quanto no campo. Além disso, como um substituto para a experimentação, ou como uma forma de estender os limites da experimentação, os economistas têm se tornado nos últimos anos entusiastas das chamadas técnicas de “variáveis instrumentais”. Por exemplo, é pouco provável que apenas examinar a correlação entre o crescimento econômico e o financiamento voltado ao desenvolvimento, mesmo controlando para outros fatores conhecidos como influenciadores do crescimento econômico, revele a influência causal do financiamento sobre o crescimento, pois o financiamento pode depender reciprocamente do crescimento e também de muitos fatores difíceis de medir que também influenciam o crescimento. Esses problemas podem ser, até certo ponto, contornados se os economistas puderem encontrar uma variável “instrumental” x da qual depende o financiamento que influencia o crescimento (se é que depende) apenas por sua influência sobre o financiamento e que é probabilisticamente independente de todos os outros determinantes do crescimento. Nesse caso, pode-se usar o efeito de x no crescimento para estimar o efeito do financiamento sobre o crescimento. Técnicas de variáveis instrumentais, experimentação de políticas e dependência de “experimentos naturais” se generalizaram, embora tragam novos problemas extrapolando resultados experimentais para a população alvo (DEATON, 2010; CARTWRIGHT; HARDIE, 2013).

2.6 Estrutura e estratégia da economia: paradigmas e programas de pesquisa

Na esteira do trabalho de (1970) e Lakatos (1970), os filósofos estão muito mais conscientes e interessados nas estruturas teóricas maiores que unificam e orientam a pesquisa dentro de tradições de pesquisa particulares. Como muitos projetos teóricos ou abordagens em economia são sistematicamente unificados, eles colocam questões sobre o que guia a pesquisa, e muitos economistas têm aplicado o trabalho de Kuhn ou Lakatos na tentativa de lançar luz sobre a estrutura geral da economia (BAUMBERG, 1977; BLAUG, 1976; DE MARCHI; BLAUG, 1991; BRONFENBRENNER, 1971; COATS, 1969; DILLARD, 1978; HANDS, 1985b; HAUSMAN, 1992, cap. 6; HUTCHISON, 1978; LATSIS, 1976; JALLADEAU, 1978; KUNIN; WEAVER, 1971; STANFIELD, 1974; WEINTRAUB, 1985; WORLAND, 1972). O sucesso dessas aplicações é controverso, mas a comparação da estrutura da economia com o esquema de Kuhn e Lakatos serviu para destacar características

distintivas da economia e pode ter contribuído para algumas das mudanças que a economia sofreu. Por exemplo, perguntar em que consiste a “heurística positiva” da economia *mainstream* permite ver que os modelos teóricos *mainstream* normalmente tentam demonstrar que um equilíbrio econômico será obtido e, portanto, que os modelos *mainstream* foram unificados em mais do que apenas suas suposições comuns. Como o sucesso dos projetos de pesquisa em economia é controverso, compreender sua estrutura e estratégia global ajudou a esclarecer seus inconvenientes e suas vantagens.

3. Inexatidão, cláusulas *ceteris paribus*, tendências, “pressupostos irrealistas” e modelos

Como mencionado na seção anterior, a questão metodológica mais importante da economia envolve a considerável simplificação, idealização e abstração que caracteriza a teoria econômica e as consequentes dúvidas que essas características da economia suscitam sobre se a economia está bem fundamentada. Alegações como “os agentes preferem pacotes de mercadorias maiores a pacotes de mercadorias menores” levantam sérias questões, pois, se interpretadas como generalizações universais, são falsas; e a filosofia da ciência tem tradicionalmente suposto que a ciência se dedica à descoberta de leis genuínas – ou seja, verdadeiras generalizações universais. Ainda que seja falso que todos sempre prefiram pacotes de mercadorias maiores a pacotes menores, a generalização parece informativa e útil. Pode uma ciência se apoiar em falsas generalizações? Se essas afirmações não são generalizações universais, qual é a sua forma lógica? E como podem ser testadas, e confirmadas ou rejeitadas, as afirmações que assim parecem ser falsas ou aproximações? Estes problemas têm atormentado economistas e metodologistas econômicos desde as primeiras reflexões metodológicas até os dias de hoje.

3.1 A economia clássica e o método *a priori*

As primeiras reflexões ampliadas sobre metodologia econômica aparecem no trabalho de Nassau Senior (1836) e John Stuart Mill (1836). Seus ensaios devem ser entendidos no contexto tanto da teoria econômica quanto da filosofia da ciência de sua época. Assim como a economia de Smith (à qual ela devia muito) e a

economia moderna, a economia “clássica” de meados do século 19 atribuiu regularidades econômicas às escolhas dos indivíduos que enfrentavam restrições sociais e naturais. Mas, em comparação a Smith, foi colocada mais confiança em modelos severamente simplificados. Os *Principles of Political Economy* [Princípios de Economia Política], de David Ricardo (1817), traça um retrato em que os salários acima do nível de subsistência levam ao aumento da população, que por sua vez requer uma agricultura mais intensiva ou o cultivo de terras inferiores. A extensão do cultivo leva a menores lucros e aluguéis mais altos; e toda a história do desenvolvimento econômico leva a um estado estacionário sombrio em que os lucros são muito baixos para comandar qualquer investimento líquido, os salários deslizam de volta aos níveis de subsistência, e apenas os proprietários são ricos.

Felizmente para o mundo, mas infelizmente para os teóricos econômicos de meados do século XIX, os dados contradiziam consistentemente as tendências que a teoria previa (DE MARCHI, 1970). No entanto, a teoria continuou a se sustentar por mais de meio século, e os dados consistentemente desfavoráveis foram explicados como sendo devido a várias “causas intervenientes”. Consequentemente, não é surpreendente que os relatos de Senior e Mill sobre o método de economia enfatizem a autonomia relativa da teoria.

Mill distingue dois tipos principais de métodos indutivos. O método *a posteriori* é um método de experiência direta. Na sua visão, ele só é adequado para fenômenos nos quais poucos fatores causais estão operando ou nos quais controles experimentais são possíveis. Os famosos métodos de indução de Mill fornecem uma articulação do método *a posteriori*. Em seu método de diferença, por exemplo, fixa-se cada fator causal exceto um e verifica-se se o efeito deixa de se obter quando esse fator é removido. O objetivo é identificar as leis causais sem exceção.

Mill afirma que métodos indutivos diretos não podem ser usados para estudar fenômenos nos quais muitos fatores causais estão em jogo. Se, por exemplo, tentarmos investigar se as tarifas aumentam ou impedem a prosperidade, comparando a prosperidade de nações com tarifas elevadas e de nações sem tarifas elevadas, os resultados não serão informativos, pois a prosperidade depende de tantos outros fatores causais. Portanto, argumenta Mill, é preciso, em vez disso, empregar o método *a priori*. Apesar de seu nome, este também é um método indutivo. Entretanto, ao contrário do método *a posteriori*, o método *a priori* é um método indutivo **indireto**. Os cientistas primeiro determinam as leis que regem os fatores causais individuais nos domínios em que os métodos de indução de Mill são aplicáveis. Depois de determinar as leis das causas individuais, eles investigam suas consequências

combinadas de forma dedutiva. Finalmente, há um papel para “verificação” das consequências combinadas, mas devido às complicações causais, este teste tem comparativamente pouco peso. O teste das conclusões serve apenas como verificação das deduções do cientista e como indicador da existência de causas intervenientes significativas, que os cientistas ainda não contabilizaram.

Mill dá o exemplo da ciência das marés. Os físicos determinaram a lei da gravitação através do estudo do movimento planetário, no qual a gravidade é o único fator causal significativo. Eles então desenvolvem a teoria das marés a partir dessa lei e informações relativas às posições e movimentos da lua e do sol. As implicações da teoria serão inexatas e às vezes muito equivocadas, pois muitos fatores causais subsidiários influenciam as marés. Testar teorias de marés pode desvendar erros nas deduções feitas pelos físicos, e pode desvendar evidências a respeito do papel dos fatores subsidiários. Mas devido à complexidade causal, tais testes pouco fazem para confirmar ou rejeitar a lei da gravitação, que já foi estabelecida. Embora Mill não use frequentemente a linguagem de “*ceteris paribus*”, sua visão de que os princípios ou “leis” da economia se sustentam na ausência de “interferências” ou “causas intervenientes” dá conta de como os princípios da economia podem ser verdadeiros *ceteris paribus* (HAUSMAN, 1992, cap. 8, 12).

Como a teoria econômica inclui apenas as causas mais importantes e necessariamente ignora as causas menores, suas alegações, como as que dizem respeito às marés, são inexatas. Suas previsões serão imprecisas, e às vezes distantes. Mill sustenta que é possível, no entanto, desenvolver e confirmar a teoria econômica ao estudar em domínios mais simples as leis que regem os principais fatores causais e então deduzir suas consequências em circunstâncias mais complexas. Por exemplo, os dados estatísticos são ambíguos quanto à relação entre salário mínimo e desemprego dos trabalhadores não qualificados; e como o salário mínimo nunca foi extremamente elevado, não há dados sobre o que seria o desemprego nessas circunstâncias. Por outro lado, a experiência cotidiana ensina aos economistas que as empresas podem escolher entre processos mais ou menos intensivos em mão-de-obra e que um salário mínimo elevado tornará mais caros os processos mais intensivos em mão-de-obra. Partindo do princípio de que as empresas tentam manter seus custos baixos, os economistas têm boas (embora não conclusivas) razões para acreditar que um salário mínimo elevado aumentará o desemprego.

Ao defender uma visão da economia como tal, inexata e empregando o método *a priori*, Mill pensou ser capaz de conciliar seu empirismo e seu compromisso com a economia de Ricardo. Embora a visão de Mill sobre a metodologia econômica

tenha sido desafiada mais tarde no século XIX por economistas que acreditavam que a teoria estava muito distante das contingências da política e da história (ROSCHER, 1874; SCHMOLLER, 1888, 1898), a visão metodológica de Mill dominou o *mainstream* da teoria econômica durante um século (*vide* CAIRNES, 1875). A visão de Mill sobreviveu à chamada revolução neoclássica na economia, a partir da década de 1870, e é claramente discernível nos mais importantes tratados metodológicos relativos à economia neoclássica, como *The Scope and Method of Political Economy* [O Escopo e o Método da Economia Política] (1891), de John Neville Keynes, ou *An Essay on the Nature and Significance of Economic Science* [Um Ensaio sobre a Natureza e o Significado da Ciência da Economia] (1932), de Lionel Robbins. Hausman (1992) argumenta que a prática metodológica atual se assemelha muito à metodologia de Mill, apesar de poucos economistas a defenderem explicitamente.

Embora esta forma de interpretar Mill e a metodologia da economia seja coerente e conforme a uma filosofia empírica da ciência antiquada que encontra a força nomológica das generalizações em sua universalidade, ela não é fiel à forma como os economistas veem suas teorias. Mais do que considerar generalizações como a ganância como leis universais com qualificações *ceteris paribus* implícitas em seus antecedentes, os economistas são muito mais propensos a considerar essas generalizações como “tendências” que continuam a operar mesmo quando derrotadas por interferências e que precisam ser estudadas separadamente (WOODWARD, 2003). Até mesmo Mill fala de tendências, embora sem conciliar seu discurso de tendências com seu empirismo. Se deixarmos de lado as dúvidas metafísicas sobre tendências e contrafactuais, a forma mais natural de ver a teorização econômica é como a investigação contrafactual de combinações de tendências. Como confirma a discussão abaixo sobre modelos, tais visões são simpáticas aos economistas e intrigantes aos filósofos com escrúpulos empiristas.

Conceptualizar a investigação econômica como o estudo de modelos e tendências, parece mudar os termos dos problemas colocados pela inexistência em vez de oferecer uma solução. Julian Reiss, de fato, redescobriu o problema em um influente ensaio, “The Explanation Paradox” [O Paradoxo da Explicação] (2013), no qual ele argumenta que as três propostas a seguir são inconsistentes: (1) Os modelos econômicos são falsos. (2) Os modelos econômicos são explicativos. (3) Explicação requer verdade. A formulação é um pouco obscura, pois modelos não são frases isoladas ou proposições que podem ser verdadeiras ou falsas, mas deve ficar claro que o paradoxo putativo de Reiss é uma reformulação do problema colocado pela inexistência das teorias ou modelos econômicos.

3.2 Friedman e a defesa dos “pressupostos irrealistas”

Embora alguns filósofos contemporâneos tenham argumentado que o método de Mill *a priori* é amplamente defensável (BHASKAR, 1975; CARTWRIGHT, 1989; HAUSMAN, 1992), em meados do século XX, a visão de Mill apareceu a muitos economistas defasada da compreensão da filosofia contemporânea da ciência por eles compartilhada. Sem estudar cuidadosamente o texto de Mill, era fácil para os economistas entenderem mal sua terminologia e considerarem seu método *a priori*, em oposição ao empirismo. Outros levaram a sério a visão de Mill de que os princípios básicos da economia deveriam ser empiricamente estabelecidos e encontraram evidências para lançar dúvidas sobre alguns dos princípios básicos, particularmente a visão de que as empresas tentam maximizar os lucros (HALL; HITCH, 1938; LESTER, 1946, 1947). Metodólogos bem informados sobre os desenvolvimentos contemporâneos da filosofia da ciência, como Terence Hutchison (1938), denunciaram a “teoria pura” da economia como não-científica.

Os economistas filosoficamente reflexivos propuseram várias formas de substituir a visão antiquada de Mill por uma metodologia mais atualizada que continuaria a justificar grande parte da prática atual (*vide* MACHLUP, 1955, 1960; KOOPMANS, 1957). De longe o mais influente desses esforços foi o ensaio de Milton Friedman, de 1953, “The Methodology of Positive Economics” [A Metodologia da Economia Positiva]. Este ensaio tem tido uma enorme influência, muito mais do que qualquer outro trabalho sobre metodologia.

Friedman começa seu ensaio distinguindo de forma convencional entre economia positiva e normativa e conjecturando que as disputas políticas são tipicamente disputas reais sobre as consequências das alternativas e podem, portanto, ser resolvidas pelo progresso da economia positiva. Passando à economia positiva, Friedman afirma (sem argumentos) que a previsão correta dos fenômenos ainda não observados é o objetivo final de todas as ciências positivas. Ele tem uma visão prática da ciência e encontra o valor da ciência nas previsões que orientarão a política.

Como é difícil o bastante e muitas vezes impossível realizar experimentos e como os fenômenos descontrolados observados pelos economistas são de difícil interpretação (devido à mesma complexidade causal que incomodava Mill), é difícil julgar se uma determinada teoria é ou não uma boa base para previsões. Tendências não são leis universais. Uma afirmação como “as empresas tentam maximizar os lucros” será “irrealista”, no sentido de que não é uma verdadeira generalização universal. Embora não nesses termos, Friedman se opõe a críticas de tendências

que, na verdade, reclamam que são meras tendências, e não leis universais. Se sua crítica parasse por aí, seria sensata, embora evitasse os problemas de compreensão e de avaliação das reivindicações sobre tendências.

Friedman, entretanto, tira uma conclusão muito mais radical. Em sua terminologia, o erro dos economistas que criticam afirmações como “as empresas tentam maximizar os lucros” está mais na tentativa de testar teorias pelo “realismo” de suas “suposições” do que pela exatidão de suas previsões. Ele sustenta que o realismo das suposições de uma teoria é irrelevante para o seu valor preditivo. Não importa se a suposição de que as empresas maximizam os lucros é realista. As teorias devem ser avaliadas exclusivamente em termos da precisão de suas predições. O que importa é exclusivamente se a teoria da empresa faz previsões corretas e significativas.

Como os críticos têm apontado (e quase todos os comentaristas têm sido críticos), Friedman se refere a várias coisas diferentes como “suposições” de uma teoria e significa várias coisas diferentes ao falar de suposições como “irrealistas” (BRUNNER, 1969). Uma vez que Friedman tem como objeto de sua crítica aqueles que investigam empiricamente se as empresas de fato tentam maximizar os lucros, ele deve tomar “suposições” para incluir generalizações econômicas centrais, como “as empresas tentam maximizar os lucros”, e por “irrealistas”, ele deve significar, entre outras coisas, “falsas”. Ao argumentar que é um erro avaliar as teorias em termos de realismo de suposições, Friedman está argumentando pelo menos que é um erro avaliar as teorias investigando se suas generalizações centrais são verdadeiras ou falsas.

Parece que esta interpretação tornaria a visão de Friedman inconsistente, pois ao testar se as empresas tentam maximizar seus lucros, está-se verificando se as previsões da teoria sobre o comportamento das empresas são verdadeiras ou falsas. Uma “suposição” como “empresas maximizam os lucros” é, em si mesma, uma previsão. Mas há uma outra rugosidade. Friedman não está preocupado com todas as predições de teorias econômicas. Na visão de Friedman, “a teoria deve ser julgada pelo seu poder de previsão exclusivamente **para a classe de fenômenos que se pretende explicar**” (1953, p. 8)¹. Os economistas estão interessados apenas em algumas das implicações das teorias econômicas. Outras previsões, como as

¹ N.T.: grifo do tradutor.

relativas aos resultados das pesquisas dos gestores, são irrelevantes para a formação de políticas públicas. O que importa é se as teorias econômicas têm sucesso em prever os fenômenos que interessam aos economistas. Em outras palavras, Friedman acredita que as teorias econômicas devem ser avaliadas em termos de suas previsões sobre preços e quantidades trocadas nos mercados. Em sua visão, o que importa é o “estrito sucesso preditivo” (HAUSMAN, 2008a) e não a adequação preditiva geral.

Assim, Friedman permite que os economistas ignorem os resultados inquietantes das pesquisas, ou o fato de as pessoas nem sempre preferirem pacotes maiores de produtos a pacotes menores de produtos. Os economistas também não precisam se preocupar se há uma tendência a preferir mais produtos a menos. Eles não precisam se preocupar que alguns de seus modelos supõem extravagantemente que todos os agentes conhecem os preços presentes e futuros de todos os produtos em todos os mercados. O que importa é se as previsões sobre os fenômenos de mercado se mostram corretas. E como os resultados anômalos do mercado podem se dever a qualquer número de fatores causais descontrolados, enquanto os experimentos são difíceis de serem realizados, ocorre que os economistas não precisam se preocupar com encontrar quaisquer evidências que falsifiquem fortemente a teoria fundamental. Modelos detalhados podem ser confirmados ou não, mas a teoria fundamental é segura. Desta forma, pode-se entender como a metodologia de Friedman, que parece justificar a visão eclética e pragmática de que os economistas devem usar qualquer modelo que pareça “funcionar”, por mais absurda ou irracional que seja sua suposição, tem sido empregada em serviço de uma ortodoxia teórica rígida. Para outras discussões sobre o ensaio de Friedman, confira Bear e Orr (1969), Boland (1979), Hammond (1992), Hirsch e de Marchi (1990), Mäki (1990a), Melitz (1963), Rotwein (1959) e Samuelson (1963).

Nas últimas duas décadas houve uma onda de experimentação em economia, e as visões metodológicas de Friedman provavelmente não atraem a mesma quase unanimidade que costumavam atrair. Mas elas ainda têm enorme influência e ainda servem como forma de evitar questões incômodas sobre simplificações, idealizações e abstrações em economia, em vez de responder a elas.

3.3 Modelos

Há um século, economistas falavam de seu trabalho em termos de “princípios”, “leis” e “teorias”. Essa linguagem não desapareceu por completo: os economistas ainda falam da “teoria dos jogos”, da “teoria da escolha do consumidor” ou da “lei da demanda”. Mas hoje a ferramenta ou forma intelectual padrão em economia é um “modelo”. Os econometristas falam de modelos e estruturas. Os economistas estão mais à vontade para descrever os axiomas relativos à escolha racional como constituindo um modelo de escolha racional do que como delineando uma teoria de escolha racional. Muitos dos mais destacados comentaristas sobre modelos os consideram como mundos ficcionais, cujo estudo informa nossa compreensão dos fenômenos reais (FRIGG, 2010). “Criar modelos é ‘fazer mundo’”. (MORGAN, 2012, p. 95, p. 405). Em sua perspectiva, os economistas são capazes de investigar como os fatores causais operariam na ausência de interferências, construindo modelos – isto é, economias fictícias – nas quais as interferências estão ausentes. Uskali Mäki sustenta que “Modelos são experimentos. Os experimentos são modelos” (2005). Dani Rodrik (2015) argumenta que economia consiste em uma coleção de modelos, e que fazer economia consiste em selecionar ou customizar um modelo dessa coleção. A ubiquidade de falar de modelos é apenas uma mudança de terminologia, ou a preocupação com modelos (que não é de forma alguma exclusiva da economia) sinaliza uma mudança metodológica? O que são modelos? Estas questões foram discutidas por Cartwright (1989, 1999), Godfrey Smith (2006), Grüne-Yanoff (2009), Hausman (1992, 2015a), Kuorikoski e Lehtinen (2009), Mäki (1991, 2005, 2009a, 2009b), Morgan (2001, 2004, 2012), Morgan e Morrison (1999), Rappaport (1998), Sugden (2000, 2009), Weisberg (2007) e Lehtinen, Kuorikoski e Ylikoski (2012).

A visão dos modelos para os quais os economistas são mais atraídos é filosoficamente problemática, pois aparentemente está comprometida com a existência de entidades ficcionais cujas propriedades e propensões causais os economistas podem investigar. Em experimentos, sejam realizados em laboratório ou no campo, os pesquisadores interagem causalmente com sujeitos experimentais de carne e osso, e o resultado pode contradizer as previsões do economista. Ao investigar um modelo, em contraste, o economista “interage” com entidades fictícias, que nada mais são do que seus próprios pensamentos, e as implicações lógicas dos axiomas que definem o modelo nunca são desapontadas. Isto não quer dizer que a investigação lógica de modelos nunca resulte em surpresas. O ser humano não é logicamente

onisciente e descobrir as implicações de um conjunto de axiomas pode ser uma tarefa árdua. Mas é uma tarefa diferente de realizar uma experiência em laboratório ou no campo, e a ontologia dos “mundos” que os economistas supostamente “criam” e depois estudam é profundamente intrigante. Embora menos fiel à prática econômica, é muito mais inteligível filosoficamente considerar os modelos como predicados ou como definições de predicados (HAUSMAN, 1992). Por exemplo, quando economistas escrevem um modelo de uma empresa com um único output e apenas dois inputs, eles estão definindo um conceito que podem usar para descrever empresas reais.

4. Abordagens influentes à metodologia econômica

O último meio século testemunhou o surgimento de uma grande literatura dedicada à metodologia econômica. Essa literatura explora muitas abordagens metodológicas e aplica suas conclusões a muitas escolas e ramos da economia. Grande parte da literatura tem se concentrado na teoria fundamental da economia geral – a teoria dos equilíbrios resultantes de uma escolha individual racional e limitada – mas a enorme importância da macroeconomia na determinação das respostas adequadas à grande recessão iniciada em 2008, aliada ao rápido aumento do papel das indagações empíricas e experimentais no trabalho cotidiano dos economistas, tem visto ecos em indagações metodológicas (BACKHOUSE, 2010). Desde 1985, existe uma revista *Economics and Philosophy* [Economia e Filosofia] dedicada especificamente à filosofia da economia, e desde 1994 existe também o *Journal of Economic Methodology* [Revista de Metodologia Econômica]. Nesta seção serão exemplificadas algumas das abordagens metodológicas das últimas duas décadas.

4.1 Abordagens popperianas

A filosofia da ciência de tem sido influente entre os economistas, assim como entre outros cientistas. Popper defende o que ele chama de uma metodologia falsificadora (1968; 1969). Os cientistas devem formular teorias que sejam “logicamente falsificáveis” – ou seja, inconsistentes com alguns possíveis resultados de observação. “Todos os corvos são negros” é logicamente falsificável; é inconsistente com (e seria falsificado por) um relatório de observação de um corvo vermelho (alegações probabilísticas obviamente não são neste sentido falsificáveis). Popper insiste na

falsificabilidade com o argumento de que alegações não falsificáveis que não descartam qualquer observação não são informativas. Elas não fornecem nenhuma orientação sobre o que esperar, e não há nada a ser aprendido ao testá-las. Segundo, Popper sustenta que os cientistas devem submeter as teorias a duros testes e devem estar dispostos a rejeitá-las quando reprovarem nos testes. Terceiro, os cientistas devem considerar as teorias como conjecturas, na melhor das hipóteses, interessantes. Passar num teste não confirma uma teoria nem dá aos cientistas motivos para acreditarem nela. Só justifica, por um lado, continuar a empregar a hipótese (por ainda não ter sido falsificada) e, por outro, dedicar mais esforços para tentar falsificá-la (por ter sobrevivido até aqui aos testes). Popper tem defendido o que ele chama de “lógica situacional” (que é basicamente a teoria da escolha racional) como o método correto para as ciências sociais (1967, 1976). Parece haver tensões sérias entre o falsificacionismo de Popper e sua defesa da lógica situacional, e sua discussão da lógica situacional não tem sido tão influente quanto seu falsificacionismo. Para a discussão de como a lógica situacional se aplica à economia, confira Hands (1985a).

Dado o falsificacionismo de Popper, parece haver pouca esperança de entender como as simplificações extremas podem ser legítimas ou como a prática econômica atual pode ser cientificamente respeitável. As teorias e os modelos econômicos são quase todos inalteráveis e, se fossem, a aceitação generalizada das visões metodológicas de Friedman asseguraria que eles não fossem submetidos a testes sérios. Quando os modelos aparentemente falham nos testes, eles raramente são repudiados. Os economistas concluem, ao invés disso, apenas que escolheram o modelo errado para a tarefa, ou que houve causas intervenientes. Os modelos econômicos, que não foram bem testados, são muitas vezes tomados como guias bem estabelecidos de políticas, ao invés de meras conjecturas. Os críticos da economia neoclássica fizeram essas críticas (EICHNER, 1983), mas a maioria dos que abraçaram a filosofia da ciência de Popper não repudiaram a economia *mainstream* e não foram duramente críticos de seus praticantes.

Mark Blaug (1992) e Terence Hutchison (1938, 1977, 1978, 2000), que são os mais proeminentes metodologistas popperianos, criticam características particulares da economia, e ambos apelam para mais testes e uma atitude mais crítica. Por exemplo, Blaug elogia Gary Becker (1976) por sua recusa em explicar diferenças nas escolhas por diferenças nas preferências, mas critica-o por não conseguir continuar e testar severamente suas teorias (1980a, cap. 14). No entanto, Blaug e Hutchison subestimam o radicalismo das opiniões de Popper e levam sua mensagem

a ser pouco mais do que a de que os cientistas devem ser críticos e preocupados em testar suas teorias.

As críticas de Blaug e Hutchison têm sido por vezes desafiadas com o argumento de que as teorias econômicas **não podem** ser testadas, devido às suas cláusulas *ceteris paribus* e às muitas suposições subsidiárias necessárias para derivar implicações testáveis (CALDWELL, 1984). Mas esta resposta ignora a insistência de Popper de que o teste requer decisões metodológicas para não atribuir falhas de previsão a erros em suposições subsidiárias ou a “interferências”. Para ver a filosofia de Popper e sua aplicabilidade à economia, confira de Marchi (1988), Caldwell (1991), Boland (1982; 1989; 1992; 1997), e Boylan e O’Gorman (2007), Backhouse (2009) e Thomas (2017).

Aplicar literalmente o ponto de vista de Popper sobre falsificação seria destrutivo. Não só a economia neoclássica, mas todas as teorias econômicas significativas seriam condenadas como não científicas, e não haveria como distinguir as teorias econômicas uma das outras. Um grande problema com uma leitura ingênua dos pontos de vista de Popper é que não se pode derivar implicações testáveis das teorias por si só. Para derivar implicações testáveis, também são necessárias suposições subsidiárias sobre distribuição de probabilidade, dispositivos de medição, **aproximações** de variáveis não mensuráveis, ausência de interferências, e assim por diante. Este é o chamado “problema Duhem-Quine” (DUHEM, 1906; QUINE, 1953; CROSS, 1982). Estes problemas surgem geralmente, e Popper propõe que eles sejam resolvidos por uma decisão metodológica de considerar o fracasso da implicação dedutível testável como sendo um fracasso da teoria. Mas em economia os pressupostos subsidiários são duvidosos e em muitos casos conhecidos como falsos. Tomar a decisão metodológica que Popper exige não é razoável e levaria a rejeitar todas as teorias econômicas.

Imre Lakatos (1970), que foi durante a maior parte de sua carreira filosófica um seguidor de Popper, oferece uma solução amplamente popperiana para este problema. Lakatos insiste que os testes são sempre comparativos. Quando as teorias enfrentam dificuldades empíricas, como sempre fazem, tenta-se modificá-las. As modificações cientificamente aceitáveis (na terminologia de Lakatos “teoricamente progressivas”) devem ter sempre algumas implicações adicionais testáveis; caso contrário, são puramente *ad hoc*. Se algumas das novas previsões forem confirmadas, então a modificação é “empiricamente progressiva”, e há razões para rejeitar a teoria não modificada e empregar a nova teoria, independentemente de quão mal sucedida, em geral, qualquer uma das teorias possa ser. Embora o

progresso possa ser difícil de ser alcançado, a visão de Lakatos não tem as mesmas implicações destrutivas que a de Popper. Lakatos parece resolver o problema de como avaliar a teoria econômica geral argumentando que o que importa é o progresso empírico ou retrocesso, e não o sucesso ou fracasso empírico. A visão de Lakatos tem sido assim mais atrativa para os metodologistas econômicos do que a de Popper.

Desenvolvendo a noção de um “paradigma” de Thomas Kuhn (1970), e algumas dicas de Popper, Lakatos também apresentou uma visão da estrutura teórica global de empreendimentos teóricos inteiros, que ele chamou de “programas de pesquisa científica”. Lakatos enfatizou que existe um “núcleo duro” de proposições teóricas básicas que definem um programa de pesquisa e que não devem ser questionadas dentro do programa de pesquisa. Além disso, os membros de um programa de pesquisa aceitam um corpo comum de heurísticas que os orientam na articulação e modificação de teorias específicas. Essas visões também têm sido atraentes para os metodologistas econômicos, uma vez que o desenvolvimento teórico em economia é fortemente limitado e que a economia parece, à primeira vista, ter um “núcleo duro”. O fato de os economistas não desistirem de postulados teóricos básicos que parecem falsos pode ser explicado e justificado por considerá-los como parte do “núcleo duro” do “programa de pesquisa neoclássica”.

No entanto, o ponto de vista de Lakatos não fornece um relato satisfatório de como a economia pode ser uma ciência respeitável, apesar de sua dependência de simplificações extremas. Isso porque é questionável se o desenvolvimento da teoria econômica neoclássica tem demonstrado progresso empírico. Por exemplo, a substituição da teoria da utilidade “cardinal” pela teoria da utilidade “ordinal” (*vide*) na década de 1930, que é geralmente considerada como um grande avanço, envolveu a substituição de uma teoria por outra que não tinha conteúdo empírico adicional. Além disso, apesar de sua ênfase na heurística como guia de modificação da teoria, Lakatos ainda enfatiza os testes. A ciência é para Lakatos mais orientada empiricamente do que a economia *mainstream* tem sido (HANDS, 1992). Também é duvidoso se os empreendimentos de pesquisa em economia têm “núcleos duros” (HOOVER, 1991; HAUSMAN, 1992, cap. 6). Para tentativas de aplicar a visão de Lakatos à economia, confira Latsis (1976) e Weintraub (1985). Como é evidente em de Marchi e Blaug (1991), escritores sobre metodologia econômica têm nos últimos anos se desencantado cada vez mais com a filosofia de Lakatos (BACKHOUSE, 2009).

Há um segundo grande problema com a filosofia da ciência de Popper, que atormenta também a visão de Lakatos. Ambos sustentam que não existe

confirmação empírica. Para algumas dúvidas tardias, confira o trabalho de Lakatos (1974). Popper e Lakatos sustentam que as evidências nunca fornecem razões para acreditar que as afirmações científicas são verdadeiras e ambos também negam que os resultados dos testes podem justificar a confiança em afirmações em esforços práticos ou em investigações teóricas. Não há evidência melhor para uma proposição não falsificada do que para outra. Nesta visão, alguém que questiona se há evidência suficiente para alguma proposição justificar a confiança nela em estudos teóricos ou para fins de política estaria fazendo o “erro” metodológico de supor que pode haver evidência em apoio a hipóteses. Com a notável exceção de Watkins (1984), poucos filósofos dentro da tradição popperiana enfrentaram essa desafiadora consequência.

4.2 A retórica da economia

Uma reação radical às dificuldades de se justificar a confianças em simplificações severas é negar que a economia passe em uma inspeção metodológica. Alexander Rosenberg (1992) sustenta que a economia só pode fazer previsões genéricas imprecisas e não pode progredir, porque se constrói em torno da psicologia popular, que uma teoria medíocre do comportamento humano e que (devido à irreducibilidade das noções intencionais) não pode ser melhorada. As teorias econômicas complexas têm valor científico apenas como matemática aplicada, não como teoria empírica. Como a economia não mostra o mesmo progresso consistente das ciências naturais, não se pode descartar a sugestão de Rosenberg de que a economia é um beco sem saída empírico. Mas sua visão de que não fez progressos e que não permite previsões quantitativas é difícil de aceitar. Por exemplo, os economistas contemporâneos são muito melhores a precificar opções de ações ou projetar leilões do que os economistas eram até mesmo uma geração atrás.

Outra reação, igualmente radical, mas oposta, é a de Deirdre McCloskey, que nega que existam quaisquer padrões metodológicos não triviais que a economia deva atender (1985, 1992, 1994, 2000; MCCLOSKEY E ZILIAK, 2003; ZILIAK E MCCLOSKEY, 2008). Em sua visão, os únicos critérios relevantes e significativos para avaliar as práticas e produtos de uma disciplina são aqueles aceitos pelos praticantes. Além de alguns padrões gerais, como honestidade e disposição para ouvir críticas, os únicos critérios justificáveis para qualquer conversa são os dos participantes. Assim, os economistas podem descartar as pretensões arrogantes

dos filósofos de julgar o discurso econômico. O que quer que um grupo de respeitados economistas tome para ser boa economia é automaticamente boa economia. Os padrões filosóficos de sucesso empírico são conversa fiada. Quem está interessado em entender o caráter da economia e em contribuir para o seu aperfeiçoamento, deve evitar a metodologia e estudar a “retórica” da economia – ou seja, os meios de argumentação e persuasão que conseguem o sucesso entre os economistas.

Os estudos de McCloskey sobre a retórica da economia têm sido valiosos e influentes (1985, esp. cap. 5-7; MCCLOSKEY; ZILIAK, 2003; ZILIAK; MCCLOSKEY, 2008), mas grande parte de seu trabalho durante as décadas de 1980 e 1990 consiste em críticas filosóficas à metodologia econômica e não em estudos sobre a retórica da economia. Suas críticas filosóficas são problemáticas, porque a posição esboçada no parágrafo anterior é difícil de defender e potencialmente autodestrutiva. É difícil de defender, pois os padrões epistemológicos já influenciaram a conversa dos economistas. Os padrões de sucesso preditivo que nos levam a ter dúvidas sobre economia já são padrões que muitos economistas aceitam. A única maneira de escapar dessas dúvidas é renunciar aos padrões que lhes deram origem. Mas a posição de McCloskey mina qualquer argumento de princípio a favor de uma mudança nos padrões. Além disso, como argumentou Rosenberg (1988), parece que os economistas se condenariam à irrelevância se entregassem a padrões de sucesso preditivos, pois é sobre tais padrões que são tomadas as decisões políticas.

McCloskey não quer, na verdade, excluir a possibilidade de os economistas serem às vezes persuadidos quando não deveriam ser ou não são persuadidos quando deveriam ser. Ela mesma critica o mau hábito que alguns economistas têm de conflitar significância estatística com importância econômica (1985, cap. 9; MCCLOSKEY; ZILIAK, 2003; ZILIAK; MCCLOSKEY, 2008). McCloskey tipicamente caracteriza a retórica de forma descritiva como o estudo do que de fato persuade, mas às vezes ela a caracteriza normativamente como o estudo do que deve persuadir (1985, cap. 2). E se a retórica é o estudo do que **deve** persuadir racionalmente, então é metodologia, não uma alternativa à metodologia. Não se pode conjurar questões sobre se a economia é uma ciência empírica de sucesso.

4.3 “Realismo” na metodologia econômica

O metodologista econômico tem prestado pouca atenção aos debates dentro da filosofia da ciência entre realistas e antirrealistas (VAN FRAASSEN, 1980;

BOYD, 1984; PSILLOS, 1999; NINILUOTO, 2002; CHAKRAVARTY, 2010; DICKEN, 2016), pois as teorias econômicas raramente postulam a existência de entidades ou propriedades não observáveis, além de variantes de “inobserváveis no cotidiano”, como crenças e desejos. Já os metodologistas têm debatido vigorosamente os objetivos da economia, mas aqueles que argumentam que os objetivos finais são preditivos (como Milton Friedman) o fazem por seu interesse nas políticas públicas, não porque busquem evitar ou resolver enigmas epistemológicos e semânticos que envolvem referências a propriedades inobserváveis.

No entanto, há dois importantes programas recentes de realismo na metodologia econômica. O primeiro, desenvolvido principalmente por Uskali Mäki, é dedicado a explorar as variedades de realismo implícito nas declarações metodológicas e nos empreendimentos teóricos dos economistas (*vide* MÄKI, 1990a, 1990b, 1990c, 2007; LEHTINEN, KUORIKOSKI; YLIKOSKI, 2012). O segundo, que é defendido por Tony Lawson e seus colegas de trabalho, principalmente na Universidade de Cambridge, deriva do trabalho de Roy Bhaskar, de 1975. (*vide* LAWSON, 1997, 2015; BHASKAR *et al.*, 1998; FLEETWOOD, 1999; BROWN; FLEETWOOD, 2003; ACKROYD; FLEETWOOD, 2004; EDWARDS, MAHONEY; VINCENT, 2014). Na visão de Lawson, pode-se traçar muitas das inadequações da economia *mainstream* (da qual ele é crítico) para uma preocupação insuficiente com a ontologia. Ao tentar identificar as regularidades na superfície dos fenômenos, os economistas ortodoxos estão condenados ao fracasso. Os fenômenos econômicos são de fato influenciados por grande número de diferentes fatores causais, e só se pode alcançar conhecimento científico dos mecanismos e tendências subjacentes, cujo funcionamento pode ser vislumbrado de forma intermitente e obscura nas relações observáveis. Os programas de Mäki e Lawson têm pouco a ver um com o outro, embora Mäki (como Mill, Cartwright e Hausman) compartilhe a preocupação de Lawson e Bhaskar com os mecanismos causais subjacentes. Veja, também, o verbete sobre **Realismo Científico**, no site da SEP, em inglês.

4.4 Metodologia econômica e estudos sociais da ciência

Ao longo de sua história, a economia tem sido objeto de um exame sociológico e metodológico. Muitas discussões sociológicas sobre economia, como a crítica de Marx à economia política clássica, têm se preocupado em identificar distorções ideológicas e assim criticar aspectos particulares da teoria econômica e

da política econômica. Como todo programa político encontra economistas que testemunham suas virtudes econômicas, há uma fonte inesgotável de material para tais críticas. Por exemplo, na esteira do quase colapso do sistema financeiro internacional em 2008, os economistas americanos que defendiam a austeridade eram em sua maioria republicanos, enquanto os que defendiam esforços para aumentar a demanda agregada eram em sua maioria democratas.

A influência da sociologia contemporânea da ciência e dos estudos sociais da ciência, aliada às dificuldades que os metodologistas têm tido para dar sentido e racionalizar a condução da economia, levaram a esforços para fundir economia e sociologia (GRANOVETTER, 1985; SWEDBERG, 1990, 2007), bem como a uma guinada sociológica dentro da própria reflexão metodológica. Em vez de mostrar que há boas evidências que sustentam os desenvolvimentos da teoria econômica ou que esses desenvolvimentos têm outras virtudes amplamente epistêmicas, metodólogos e historiadores, como D. Wade Hands (2001), Hands e Mirowski (1998), Philip Mirowski (1990, 2002, 2004, 2013) e E. Roy Weintraub (1991), argumentaram que essas mudanças refletem uma grande variedade de fatores não racionais, desde mudanças no financiamento da economia teórica, compromissos políticos, rivalidades pessoais, apego a metáforas ou interesses matemáticos.

Além disso, muitos dos mesmos metodologistas e dos historiadores têm argumentado que a economia não é apenas um objeto de investigação social, mas que pode ser um instrumento de investigação social da ciência. Ao estudar a estrutura de incentivo das disciplinas científicas e as forças de mercado implícitas ou explícitas que incidem sobre a pesquisa (incluindo, naturalmente, a pesquisa em economia), deveria ser possível escrever a economia da ciência e a economia da própria economia (HANDS, 1995; HULL, 1988; LEONARD, 2002; MIROWSKI; SENT, 2002).

Exatamente como este trabalho supostamente, se é que ele deve o fazer, responde às perguntas sobre o quão bem apoiadas são as reivindicações dos economistas não está claro. Embora fugindo da metodologia tradicional, a monografia de Mirowski sobre o papel da analogia física na economia (1990) é muitas vezes muito crítica em relação à economia *mainstream*. Em *Reflexão sem Regras* (2001), D. W. Hands sustenta que regras metodológicas gerais são de pouca utilidade. Ele defende uma visão naturalista da metodologia e é cético em relação às prescrições que não se baseiam em conhecimentos detalhados. Mas ele não argumenta que nenhuma regra se aplica.

4.5 Estudos de caso

O levantamento acima descrito das abordagens aos problemas fundamentais da avaliação da teoria econômica está longe de ser completo. Por exemplo, tem havido esforços substanciais para aplicar visões estruturalistas das teorias científicas (SNEED, 1971; STEGMÜLLER, 1976, 1979) à economia (STEGMÜLLER *et al.*, 1981; HAMMINGA, 1983; HANDS, 1985c; BALZER; HAMMINGA, 1989). A discussão acima documenta a diversidade e as discordâncias sobre como interpretar e avaliar as teorias econômicas. Não é surpreendente que não haja consenso entre aqueles que escrevem sobre metodologia econômica em relação à avaliação empírica geral de abordagens específicas em economia, incluindo a microeconomia *mainstream*, a macroeconomia e a econometria. Quando os profissionais não conseguem concordar, é questionável se aqueles que conhecem mais filosofia, mas menos economia, serão capazes de resolver a questão. Como os debates continuam, aqueles que refletem sobre a metodologia econômica devem ter um papel contínuo a desempenhar.

Entretanto, há muitas outras questões metodológicas mais específicas a serem abordadas, e é um sinal da maturidade da subdisciplina que uma grande e crescente porcentagem de trabalho em metodologia econômica trata de questões mais específicas. Há uma infinidade de trabalhos, como se confirmará em qualquer edição recente do *Journal of Economic Methodology* ou *Economics and Philosophy*. Algumas das questões atualmente em discussão foram mencionadas acima na Aqui está uma lista de três das muitas áreas de interesse atual:

1. Embora mais preocupada com o conteúdo da economia do que com sua metodologia, a recente explosão do trabalho em economia feminista é feita com autorreflexão metodológica e sociológica. O fato de uma porcentagem consideravelmente maior de economistas serem homens do que a de qualquer outra ciência social e, na verdade, do que a maioria das ciências naturais, levanta questões sobre se há algo de particularmente masculino na disciplina. Textos importantes são Ferber e Nelson (1993, 2003), Nelson (1995, 1996, 2001), Barker e Kuiper (2003). Desde 1995, existe uma revista, *Feminist Economics* [Economia Feminista], que reúne grande parte desse trabalho.

2. Durante as últimas décadas, a experimentação laboratorial em economia se expandiu rapidamente. A experimentação laboratorial tem muitos objetivos diferentes (*vide* ROTH, 1988) e, aparentemente, oferece a perspectiva de superar o abismo entre a teoria econômica fundamental e as evidências empíricas. Algumas

delas lançam luz sobre a forma como os compromissos metodológicos influenciam a medida em que os economistas atentam para a evidência empírica. Uma boa parte da experimentação laboratorial na economia contemporânea está a serviço da economia comportamental, que se orgulha de atender à evidência experimental relativa à estrutura e aos determinantes das escolhas individuais. Embora a economia comportamental tenha garantido uma base dentro da economia *mainstream*, ela permanece controversa substantiva e metodologicamente, e suas implicações para a economia normativa, discutidas abaixo na seção 6, são controversas.

Por exemplo, no caso das reversões de preferências, discutidas brevemente abaixo na Seção 5.1, economistas dedicaram considerável atenção aos achados experimentais e admitiram que eles refutaram princípios centrais da economia. Mas economistas geralmente não se dispuseram a prestar séria atenção às teorias propostas pelos psicólogos que previam os fenômenos antes de serem observados. A razão parece ser que essas teorias psicológicas não têm o mesmo escopo amplo dos princípios básicos da economia *mainstream* (HAUSMAN, 1992, cap. 13). Hesitação sobre neuroeconomia (CAMERER *et al.*, 2005; CAMERER, 2009; MARCHIONNI; VROMEN, 2014; RUSTICHINI, 2005, 2009; GLIMCHER; FEHR, 2013; REUTER; MONTAG, 2016; VROMEN; MARCHIONNI, 2018) também é comum. Em um ensaio extremamente influente, *The Case for Mindless Economics* [O Caso da Economia Sem Mente], Gul e Pesendorfer (2008) argumentam que os achados da economia comportamental (e neuroeconômica) são irrelevantes para a economia. Eles são, no máximo, de valor heurístico. Isto porque não se referem a escolhas de mercado e suas consequências, que são os únicos dados pertinentes. Algumas vezes Gul e Pesendorfer parecem identificar a teoria econômica com as consequências empíricas com que os economistas estão preocupados, enquanto em outros pontos eles fazem eco a Milton Friedman (*vide* seção 3.2) e negam que o “realismo” das “suposições” dos modelos econômicos importe. Eles não tratam de sofisticadas defesas do realismo em relação a estados mentais como Dietrich e List (2016). Parece-me que a resistência teórica ao engajamento com economistas comportamentais como aquele que se encontra no ensaio de Gul e Pesendorfer está se enfraquecendo. Mas é claro que os compromissos metodológicos que regem a economia teórica são muito mais complexos e específicos à economia do que as regras gerais propostas por filósofos como Popper e Lakatos.

A relevância da experimentação laboratorial permanece controversa. Economistas comportamentais são entusiastas, enquanto os teóricos mais tradicionais questionam se os resultados experimentais podem ser generalizados para contextos

não experimentais e, mais genericamente, sobre as possibilidades de aprendizado a partir de experimentos (CAPLIN; SCHOTTER, 2008). Para discussões sobre economia experimental, confira os trabalhos de Guala (2000a, 2000b, 2005), Hey (1991), Kagel e Roth (1995, 2016), Plott (1991), Smith (1991), Starmer (1999), Camerer (2003), Bardsley e Cubitt (2009), Durlauf e Blume (2009), Branas-Garza e Cabrales (2015), Fréchette e Schotter (2015), Jacquemet e L'Haridon (2018), e a edição especial, de junho de 2005, do *Journal of Economic Methodology*. A página de Al Roth, *Game Theory, Experimental Economics, and Market Design* [Teoria dos Jogos, Economia Experimental e Design de Mercado]² é também uma fonte útil. Para trabalhos recentes sobre economia comportamental veja *Journal of Behavioral Economics* [Revista de Economia Comportamental], *Review of Behavioral Economics* [Ensaios de Economia Comportamental] e *Behavioural Public Policy* [Política Pública Comportamental].

3. Na última geração, houve uma transformação radical nas atitudes dos economistas em relação à investigação causal empírica, especialmente na forma de experimentos de campo e experimentos naturais, muitas vezes empregando variáveis instrumentais. Por exemplo, cerca de dois terços dos artigos da *American Economic Review* [Revista Americana de Economia] de fevereiro de 2018 são baseados em estudos empíricos. Os títulos dos primeiros quatro artigos do índice são: *The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges* [Os efeitos da execução provisória da pena na convicção, nos crimes futuros e no emprego: evidências a partir de juízes randomicamente atribuídos], *Implications of US Tax Policy for House Prices, Rents, and Homeownership* [Implicações da política tributária dos EUA para preços de casas, aluguéis e titularidade da casa própria], *The Welfare Cost of Perceived Policy Uncertainty: Evidence from Social Security* [O custo de bem-estar resultante de incerteza percebida na política pública: evidência da seguridade social], *The Economic Consequences of Hospital Admissions* [As consequências econômicas das admissões hospitalares]. Se recuarmos vinte e cinco anos, apenas cerca de um oitavo da primeira edição da *American Economic Review* de 1993 parece confiar em quaisquer estudos empíricos. Os primeiros quatro artigos são: *Today's Task for Economists* [A tarefa dos economistas hoje], *Trigger Points and Budget Cuts: Explaining the Effects of Fiscal Austerity* [Pontos de gatilho e cortes no orçamento: explicando os

² Disponível em: <http://kuznets.fas.harvard.edu/~aroth/alroth.html>.

efeitos da austeridade fiscal], *Economic Policy, Economic Performance, and Elections* [Política Econômica, atuação econômica e eleições], *The Macroeconomics of Dr. Strangelove* [A macroeconomia do Dr. Strangelove]. Um Rip Van Winkle, que tinha ido dormir em 1983, lendo as principais revistas de economia, ficaria desconcertado quando ele acordasse em 2018.

As experiências de campo têm sido especialmente importantes em economia do desenvolvimento, onde os resultados de vários projetos de financiamento internacional têm proporcionado, com demasiada frequência, magros benefícios. Pode-se encontrar boas introduções a este trabalho em Carpenter *et al.* (2005), Duflo e Banerjee (2011, 2017), Gugerty e Karlan (2018), Karlan e Appel (2011, 2016), Kremer e Glennerster (2011), List e Samek (2018), e Mullainathan e Shafir (2013). Confira, também, o *Poverty Action Lab*³. Embora os experimentos de campo pareçam ser perguntas difíceis que estabelecem o que funciona e o que não funciona, as coisas não são tão simples (DEATON, 2010; CARTWRIGHT; HARDIE, 2013). Sem o conhecimento dos mecanismos, é muito fácil para uma intervenção que funciona esplendidamente bem em um momento e lugar específico falhar de forma abismal quando tentada em outro lugar. A investigação sem teoria, mesmo quando metodologicamente sofisticada, tem limites severos como tática de aquisição de conhecimento.

A volta empírica na economia também teve o efeito de aumentar a importância da história econômica. Com algum engenho, especialmente na identificação de possíveis variáveis instrumentais, a história está repleta de “experimentos naturais”. Por exemplo (HAUSMAN, 2016), em 1936, o Congresso americano votou a favor do pagamento de pensões aos veteranos da Primeira Guerra Mundial oito anos antes do que deveriam ser pagas. Como as percentagens de veteranos diferem entre os estados, Hausman pôde usar os diferentes desempenhos econômicos dos estados para estimar os efeitos do estímulo econômico que as pensões proporcionam. Embora menos decisivo do que as experiências controladas aleatórias (muitas vezes impossíveis de serem realizadas), o exame de episódios históricos como este fornecem evidências significativas no que diz respeito às hipóteses econômicas.

³ Disponível em: <https://www.povertyactionlab.org/>. Acesso em: 02 jan. 2022.

5. Teoria da escolha racional

Na medida em que a economia explica e prevê fenômenos como consequências de escolhas individuais, que são elas próprias explicadas em termos de alegadas razões, ela deve retratar os agentes como, em certa medida, racionais. A racionalidade, como as razões, envolve a avaliação, e assim como se pode avaliar a racionalidade das escolhas individuais, também se pode avaliar a racionalidade das escolhas sociais e examinar como elas estão e devem estar relacionadas com as preferências e julgamentos dos indivíduos. Além disso, há questões intrincadas relativas à racionalidade em situações estratégicas em que os resultados dependem das escolhas de múltiplos indivíduos. Como a racionalidade é um conceito central em ramos da filosofia como a teoria da ação, epistemologia, ética e filosofia da mente, os estudos da racionalidade frequentemente cruzam as fronteiras entre economia e filosofia.

5.1 Racionalidade individual

A essência da teoria da racionalidade discutida acima trata de um agente (ordenações de estados de coisas) como racionais se forem completas e transitivas e trata a escolha do agente como racional se ele não preferir nenhuma alternativa viável àquela que ele escolhe. Tal teoria da racionalidade é claramente muito fraca, pois nada diz sobre a crença ou o que a racionalidade implica quando os agentes não sabem (com certeza) tudo o que é relevante para suas escolhas. Mas também pode ser forte demais, pois, como Isaac Levi em particular argumentou (1986), não há nada de irracional em ter preferências incompletas em situações de incerteza. Às vezes é racional suspender o julgamento e recusar-se a classificar alternativas que não são bem compreendidas. Por outro lado, a transitividade é uma condição plausível e o argumento da chamada “bomba de dinheiro” demonstra que se as preferências de alguém são intransitivas e se ele está disposto a fazer trocas, então pode ser explorado. Suponha que um agente **A** prefira **X** a **Y**, **Y** a **Z** e **Z** a **X**, e que **A** pague uma pequena quantia de dinheiro \$**P** para trocar **Y** por **X**, **Z** por **Y**, e **X** por **Z**. Isso significa que, começando por **Z**, **A** pagará \$**P** por **Y**, depois \$**P** novamente por **X**, depois \$**P** novamente por **Z** e assim por diante. Os agentes não são tão estúpidos assim. Eles vão se recusar a negociar ou ajustar suas preferências para eliminar a intransitividade (*vide* SCHICK, 1986).

Por outro lado, há considerável evidência experimental de que as preferências das pessoas de fato não são transitivas. Tais evidências não estabelecem que a transitividade não é um requisito de racionalidade; podem mostrar, ao invés disso, que as pessoas às vezes são irracionais. No caso das chamadas “inversões de preferências”, por exemplo, parece plausível que as pessoas de fato façam escolhas irracionais (LICHTENSTEIN; SLOVIC, 1971; TVERSKY; THALER, 1990). A evidência de persistentes violações da transitividade é inquietante, pois os padrões de racionalidade não devem ser impossivelmente altos.

Uma outra dificuldade com os argumentos principais da teoria da racionalidade diz respeito à individuação dos objetos de preferência ou escolha. Considere, por exemplo, os dados de jogos de ultimato com várias etapas. Suponha que **A** pode propor qualquer divisão de \$10 entre **A** e **B**. **B** pode aceitar ou rejeitar a proposta de **A**. Se **B** rejeitar a proposta, então a quantidade de dinheiro cai para \$5, e **B** consegue oferecer uma divisão dos \$5 que **A** pode aceitar ou rejeitar. Se **A** rejeitar a oferta de **B**, então ambos os jogadores não recebem nada. Suponha que **A** propõe dividir o dinheiro com \$7 para **A** e \$3 para **B**. **B** recusa e oferece para dividir os \$5 igualmente, com \$2,50 para cada um. Comportamentos como estes são, na verdade, comuns (Ochs e Roth, 1989, p. 362). Assumindo que **B** prefere mais dinheiro a menos, essas escolhas parecem ser uma violação da transitividade. **B** prefere \$3 a \$2,50, ainda assim diminui \$3 para certo por \$2,50 (com alguma pequena chance de **A** recusar e **B** não receber nada). Mas os objetos de escolha não são apenas quantidades de dinheiro. **B** está recusando \$3 como parte de um “acordo forçado” em favor de \$2,50 como parte de um acordo justo. Se os objetos de escolha são definidos desta forma, não há falha de transitividade.

Esta observação plausível dá origem a um grave problema. A menos que haja restrições sobre como os objetos de escolha são individualizados, condições de racionalidade como a transitividade são vazias. A escolha de **A** por **X** sobre **Y**, **Y** sobre **Z** e **Z** sobre **X** não viola a transitividade se “**X** quando a alternativa é **Y**” não for o mesmo objeto de escolha que “**X** quando a alternativa é **Z**”. John Broome (1991) argumenta que outros princípios substantivos de racionalidade são necessários para limitar como as alternativas são individualizadas ou para exigir que os agentes sejam indiferentes entre alternativas como “**X** quando a alternativa é **Y**” e “**X** quando a alternativa é **Z**”.

Para estender a teoria da racionalidade a circunstâncias envolvendo (em que os objetos de escolha são loterias com probabilidades conhecidas) e incerteza (em que os agentes não conhecem as probabilidades ou mesmo todos os resultados

possíveis de suas escolhas), são necessários mais princípios de racionalidade, assim como simplificações técnicas controversas. Os subjetivistas Bayesianos supõem que indivíduos em circunstâncias de incerteza têm probabilidades subjetivas bem definidas (graus de crença) sobre todos os resultados e, portanto, os objetos de escolha podem ser modelados como loterias, assim como em circunstâncias de risco, embora com subjetivas, condições de risco e incerteza é a condição de independência. Diz aproximadamente que as preferências do agente racional entre duas loterias que diferem em apenas um resultado devem corresponder às suas preferências entre os diferentes resultados. Embora inicialmente plausível, a condição de independência é muito controversa. Ver Allais e Hagen (1979) e McClennen (1983, 1990).

Uma parte considerável da teoria da escolha racional diz respeito à formalização das condições de racionalidade e à investigação das suas implicações. Quando as preferências de um agente são completas e transitivas e satisfazem outra condição de continuidade, então elas podem ser representadas pela chamada função de utilidade ordinal. Isto significa que é possível definir uma função que representa as preferências de um agente para que $U(X) > U(Y)$ se e somente se o agente preferir X a Y e $U(X) = U(Y)$ se, e somente se, o agente for indiferente entre X e Y . Esta função representa meramente a ordenação de preferências. Ela não contém informações além da ordenação. Qualquer transformação que preservasse a ordenação de “ U ” representaria as preferências do agente da mesma forma.

Quando as preferências de um agente satisfazem além disso a condição de independência e algumas outras condições técnicas, então eles podem ser representados por uma função de utilidade esperada (HARSANYI, 1977b, cap. 4; HERNSTEIN e MILNOR, 1953; RAMSEY, 1926; e SAVAGE, 1972). Tal função tem duas propriedades importantes. Em primeiro lugar, a utilidade esperada de uma loteria é igual à soma das utilidades (esperadas) de seus prêmios ponderadas por suas probabilidades. Em segundo lugar, as funções de utilidade esperada são únicas – exceto por transformações afins positivas. Isto significa que se U e V são ambas funções de utilidade esperada representando as preferências de um agente, então para todos os objetos de preferência, X , $V(X)$ deve ser igual a $aU(X) + b$, onde a e b são números reais e a é positivo. Além disso, os axiomas de racionalidade implicam que os graus de crença do agente irão satisfazer os axiomas do cálculo de probabilidade.

Muita controvérsia envolve a teoria da racionalidade, tendo sido executadas muitas investigações formais sobre versões mais fracas (logicamente) ou alteradas de teorias da racionalidade. Para mais discussões, ver Allais e Hagen (1979), Barberà, Hammond e Seidl (1999), Kahneman e Tversky (1979), Loomes e Sugden

(1982), Luce e Raiffa (1957), Machina (1987) e Gilboa e Schmeidler (2001).

5.2 Racionalidade coletiva e escolha social

Embora as sociedades sejam muito diferentes dos indivíduos, elas têm mecanismos para avaliar alternativas e fazer escolhas, e suas avaliações e escolhas podem ser racionais ou irracionais. Não é óbvio, porém, quais princípios de racionalidade devem reger as escolhas e avaliações da sociedade. A transitoriedade é uma condição plausível. Parece que uma sociedade que escolhe **X** quando confrontada com as alternativas **X** ou **Y**, **Y** quando diante das alternativas **Y** ou **Z** e **Z** quando diante das alternativas **X** ou **Z** ou teve uma mudança de opinião ou está escolhendo irracionalmente. No entanto, supostas irracionalidades como estas podem facilmente surgir a partir de mecanismos padrão que visam a vincular escolhas sociais e preferências individuais. Suponha que existam três indivíduos na sociedade. O indivíduo Um classifica as alternativas **X**, **Y**, **Z**. O indivíduo Dois classifica **Y**, **Z**, **X**. O indivíduo Três classifica **Z**, **X**, **Y**. Se as decisões forem tomadas entre pares de opções, por maioria de votos, **X** será escolhido do par (**X**, **Y**), **Y** será escolhido de (**Y**, **Z**), e **Z** será escolhido de (**X**, **Z**). É claro que isso é inquietante, mas os possíveis ciclos de escolhas sociais são **irracionais**?

Problemas semelhantes afetam o que se poderia chamar de coerência lógica dos julgamentos sociais (LIST; PETTIT, 2002). Suponhamos que a sociedade seja constituída por três indivíduos que fazem os seguintes julgamentos sobre a verdade ou falsidade das proposições **P** e **Q** e que o julgamento social segue a maioria.

	P	Se P , então Q	Q
Indivíduo 1	verdadeiro	verdadeiro	Verdadeiro
Indivíduo 2	falso	verdadeiro	falso
Indivíduo 3	verdadeiro	falso	falso
Sociedade	verdadeiro	verdadeiro	falso

Os julgamentos de cada um dos indivíduos são consistentes com os princípios da lógica, enquanto os julgamentos sociais os violam. Quão importante é que os julgamentos sociais sejam consistentes com os princípios da lógica?

Embora a teoria da escolha social, desta forma, se baseie em questões de racionalidade social, a maioria dos trabalhos em teoria da escolha social explora

as consequências dos princípios de racionalidade aliados a restrições explicitamente éticas. A contribuição seminal é o teorema da impossibilidade de Kenneth Arrow (1963, 1967). Arrow assume que tanto as preferências individuais quanto as preferências sociais são completas e transitivas e que o método de formação de preferências sociais (ou de fazer escolhas sociais) questiona em algum ranking de preferências sociais ou escolha social para qualquer perfil possível de preferências individuais. Além disso, Arrow impõe uma condição de unanimidade fraca: se todos preferem **X** a **Y**, então **Y** não deve ser preferido socialmente. Terceiro, ele exige que não haja ditador cujas preferências determinem preferências ou escolhas sociais, independentemente das preferências de qualquer outra pessoa. Finalmente, ele impõe a condição de que a preferência social entre **X** e **Y** deve depender de como os indivíduos ordenam **X** e **Y** e de nada mais. Arrow então provou o resultado surpreendente de que nenhum método de relacionar preferências sociais e individuais pode satisfazer todas essas condições!

Nos sessenta anos desde que Arrow escreveu sobre o tema, houve uma infinidade de trabalhos no campo da teoria da escolha social, boa parte dos quais, sem dúvida, de grande importância para a ética. Por exemplo, John Harsanyi provou que se as preferências individuais e as avaliações sociais satisfazem os axiomas da teoria da utilidade esperada (com probabilidades compartilhadas ou objetivas) e que as preferências sociais estão de acordo com as preferências individuais unânimes, então as avaliações sociais são determinadas por uma soma ponderada de utilidades individuais (1955, 1977a). Matthew Adler (2012) estendeu uma abordagem como a de Harsanyi para demonstrar que uma forma de utilitarismo ponderado, que prioriza os interesses daqueles que estão em pior situação, satisfaz de forma única uma lista mais longa de restrições racionais e éticas. Quando, ao invés disso, há discordâncias nas atribuições de probabilidade, há um resultado impossível: a condição de unanimidade implica que, para alguns perfis de preferências individuais, as avaliações sociais não satisfarão os axiomas da teoria da utilidade esperada (HAMMOND, 1983; SEIDENFELD *et al.*, 1989; MONGIN, 1995). Para mais discussões sobre a teoria da escolha social e a relevância da teoria da utilidade para a avaliação social, ver o verbete sobre a *Teoria da escolha Social*, no site da SEP, em inglês⁴, Sen (1970) e, para as recentes reavaliações, Fleurbaey (2007) e Adler (2012).

⁴ N.T.: Esse verbete também se encontra traduzido nesse volume.

5.3 Teoria dos jogos

Quando os resultados dependem do que vários agentes fazem, a melhor escolha de um agente pode depender do que os outros agentes escolhem. Embora os princípios de racionalidade que governam a escolha individual ainda se apliquem, é discutível que existem outros princípios de racionalidade que governam as expectativas das ações dos outros (e de suas expectativas em relação às suas ações e expectativas, e assim por diante). A teoria dos jogos ocupa um papel cada vez mais importante dentro da economia e também é relevante tanto para as questões relativas à racionalidade quanto para as questões relativas à ética. Para mais discussões veja os verbetes sobre *Teoria dos Jogos*, *Teoria dos Jogos e Ética** e *Teoria Evolutiva dos Jogos*, no site da SEP, em inglês.

6. Economia e ética

Como discutido acima na a maioria dos economistas faz distinção entre economia positiva e economia normativa, e a maioria argumentaria que a economia é relevante para a política principalmente devido às informações (positivas) que ela fornece sobre as consequências da política. No entanto, os mesmos economistas também oferecem seus conselhos sobre como fixar a economia, havendo todo um campo da economia normativa.

Os resultados econômicos, instituições e processos podem ser melhores ou piores de várias maneiras diferentes. Alguns resultados podem melhorar a vida das pessoas. Outros resultados podem ser menos desiguais. Outros podem restringir a liberdade individual mais severamente. Os economistas costumam avaliar os resultados exclusivamente em termos de bem-estar. Isso não significa que eles acreditem que apenas o bem-estar tem importância moral. Eles se concentram no bem-estar, porque acreditam que a economia fornece um excelente conjunto de ferramentas para abordar questões de bem-estar e porque esperam que as questões sobre o bem-estar possam ser separadas das questões sobre igualdade, liberdade ou justiça. Conforme esboçado abaixo, os economistas têm algumas coisas a dizer sobre outras dimensões da avaliação moral, mas o bem-estar é o centro das atenções. Na verdade, a economia normativa é normalmente chamada de “economia do bem-estar”.

6.1 Bem-estar

Uma questão central da filosofia moral é a de determinar quais coisas são intrinsecamente boas para os seres humanos. Essa é uma questão central, porque todas as visões morais plausíveis atribuem um lugar importante ao bem-estar. Isso é obviamente verdadeiro para o utilitarismo (que sustenta que o correto é maximizar o bem-estar total ou médio), mas mesmo as visões não utilitaristas estão preocupadas com o bem-estar – se reconhecerem a virtude da benevolência, ou se estiverem preocupadas com os interesses dos indivíduos, ou com evitar danos.

Existem muitas maneiras de pensar sobre o bem-estar, e a visão predominante entre os economistas mudou do hedonismo (que considera o bem como um estado mental como prazer ou felicidade) para a visão de que o bem-estar deve ser medido pela satisfação das preferências. Vários economistas proeminentes estão atualmente defendendo um retorno ao hedonismo, mas eles permanecem uma minoria. (*vide* BAVETTA *et al.*, 2014; FLÈCHE, 2018; DOLAN; KAHNEMAN, 2014; FREY, 2010, 2018; FREY; STUTZER, 2001; KAHNEMAN, 1999, 2000a, 2000b; KAHNEMAN; KRUEGER, 2006; KAHNEMAN; SUGDEN, 2005; KAHNEMAN; THALER, 2006; LAYARD, 2006; ORMEROD, 2008; RADCLIFF, 2013; WEIMANN; KNABE, 2015). Para críticas (DAVIES, 2015; ETZIONI, 2018; HAUSMAN, 2010). Ao contrário do hedonismo, considerar o bem-estar como satisfação de preferência implica especificar de que forma descobrir o que é bom para uma pessoa, em vez de se comprometer a qualquer visão substantiva do bem de uma pessoa. Observe que igualar o bem-estar à satisfação de preferências não é igualar o bem-estar a qualquer sentimento de satisfação. Se o bem-estar pode ser medido pela satisfação de preferências, então uma pessoa ficará melhor se o que ela prefere acontecer, independentemente de essa ocorrência deixar o agente satisfeito.

Uma vez que a economia *mainstream* atribui uma ordenação de preferência consistente a todos os agentes, e uma vez que modelos mais específicos normalmente consideram os agentes bem-informados e com interesse próprio, é fácil para os economistas aceitarem a visão de que um agente individual *A* preferirá *X* a *Y* se, e somente se, *X* for, de fato, melhor do que *Y* para *A*. Este é um ponto em que a teoria positiva se transforma em teoria normativa. Além disso, a identificação do bem-estar com a satisfação das preferências é atraente para os economistas, porque evita que surjam dúvidas sobre a justificativa do paternalismo (ao qual a maioria dos economistas se opõe fortemente).

O bem-estar e a satisfação das preferências podem coincidir porque a satisfação das preferências constitui o bem-estar ou porque as pessoas se interessam por si mesmas e julgam seus próprios interesses e, portanto, preferem o que é bom para elas. Existem muitas objeções óbvias à visão de que a satisfação das preferências constitui bem-estar. As preferências podem ser baseadas em crenças equivocadas. As pessoas podem preferir sacrificar seu próprio bem-estar por algum propósito que valorizam mais. As preferências podem refletir manipulações passadas ou distorções de influências psicológicas (ELSTER, 1983). Além disso, se a satisfação de preferências constitui bem-estar, então os formuladores de políticas podem tornar as pessoas melhores moldando seus desejos, em vez de melhorar as condições. Além disso, não parece razoável que a política social deva atender a preferências extravagantes. Em vez de responder a essas objeções e tentar defender a visão de que a satisfação das preferências constitui bem-estar, os economistas podem atenuar essas objeções ao considerar as preferências em circunstâncias em que as pessoas estão interessadas em si mesmas e bons juizes de seus interesses são apenas boas evidências do que irá promover o bem-estar (HAUSMAN; MCPHERSON, 2009; HAUSMAN, 2012). Existem algumas exceções, mais notavelmente Amartya Sen (1987a, 1987b, 1987c, 1992), mas a maioria dos economistas considera que o bem-estar coincide com a satisfação da preferência.

6.2 Eficiência

Como a identificação do bem-estar com a satisfação das preferências torna questionável se é possível fazer comparações interpessoais de bem-estar, poucos economistas defendem uma visão utilitarista da política como maximização do bem-estar total ou médio. Harsanyi é uma exceção; para outra, confira Ng (1983). Em vez disso, os economistas exploraram a possibilidade de fazer avaliações de bem-estar de processos, instituições, resultados e políticas econômicas sem fazer comparações interpessoais. Considere dois resultados econômicos **S** e **R**, e suponha que algumas pessoas prefiram **S** a **R** e que ninguém prefere **R** a **S**. Nesse caso, **S** é “Pareto superior” a **R**, ou **S** é uma “melhoria de Pareto” em relação a **R**. Sem fazer quaisquer comparações interpessoais, pode-se concluir que as preferências das pessoas são mais bem satisfeitas em **S** do que em **R**. Se não há estado de coisas que seja Pareto superior a **S**, então os economistas dizem que **S** é “Pareto ótimo” ou “Pareto eficiente”. Eficiência aqui é a eficiência com relação à satisfação de

preferências, em vez de minimizar o número de insumos necessários para produzir uma unidade de produto ou alguma outra noção técnica (LE GRAND, 1991). Se um estado de coisas não é Pareto-eficiente, então a sociedade está perdendo uma oportunidade sem custos de satisfazer melhor as preferências de algumas pessoas. Um estado de coisas Pareto-eficiente evita esse fracasso, mas não tem outras virtudes óbvias. Por exemplo, suponha que ninguém esteja saciado e as pessoas se importem apenas com a quantidade de comida que recebem. Considere duas distribuições de alimentos. No primeiro, milhões estão morrendo de fome, mas nenhum alimento é desperdiçado. No segundo, ninguém passa fome, mas algum alimento é desperdiçado. O primeiro é Pareto eficiente, enquanto o segundo não.

As noções de melhorias de Pareto e eficiência de Pareto podem parecer inúteis, porque as políticas econômicas quase sempre têm vencedores e perdedores. Economistas *mainstream*, no entanto, consideram esses conceitos úteis de duas maneiras. Primeiro, eles provaram dois teoremas relativos a propriedades de equilíbrios perfeitamente competitivos (ARROW, 1968). O primeiro teorema diz que os equilíbrios em mercados perfeitamente competitivos são ótimos de Pareto, e o segundo diz que qualquer alocação ótima de Pareto, com qualquer distribuição que formuladores de políticas possam preferir, pode ser alcançada como um equilíbrio de mercado perfeitamente competitivo, desde que se comece com apenas a correta distribuição de dotações entre os agentes econômicos. O primeiro teorema foi considerado como sustentando a visão de Adam Smith da mão invisível (ARROW; HAHN, 1971, prefácio; HAHN, 1973). Essa interpretação é problemática, porque nenhuma economia jamais esteve ou estará em equilíbrio perfeitamente competitivo. O segundo teorema fornece alguma justificativa para a divisão de trabalho normativa que economistas preferem, com uns economistas preocupados com a eficiência e outros preocupados com a justiça. A ideia é que o segundo teorema mostra que as teorias de distribuição justa são compatíveis com a dependência de mercados competitivos. Os dois teoremas fundamentais da economia do bem-estar explicam por que os economistas tradicionais, estejam eles apoiando políticas de *laissez-faire* ou intervenção governamental para remediar imperfeições de mercado, pensam em equilíbrios perfeitamente competitivos como ideais. Mas a importância dos teoremas é discutível, uma vez que os mercados reais diferem significativamente de mercados perfeitamente competitivos e, quando há várias imperfeições de mercado, a “teoria do segundo melhor” mostra que consertar algumas das imperfeições pode levar a sociedade para longe de um perfeito equilíbrio competitivo (e diminuir a eficiência e o bem-estar) em vez de em direção a um equilíbrio competitivo (LIPSEY; LANCASTER, 1956-7).

A outra maneira que os economistas encontraram para estender as noções de eficiência de Pareto leva à análise de custo-benefício, que é uma ferramenta prática para a análise de políticas (MISHAN, 1971; SUGDEN; WILLIAMS, 1978; ADLER; POSNER, 2000, 2006; BROADMAN *et al.*, 2010; BOADWAY, 2016). Suponha que **S** não seja uma melhoria de Pareto em relação a **R**. Alguns membros da sociedade seriam perdedores em uma mudança de **R** para **S**. Esses perdedores preferem **R** a **S**, mas há vencedores suficientes – pessoas suficientes que preferem **S** a **R** – que os vencedores poderiam compensar os perdedores e tornar a preferência por **S'** (**S** com compensação paga) em vez de **R** unânime. **S** é uma “melhoria potencial de Pareto” em relação a **R**. Em outros termos, a quantidade de dinheiro que os vencedores estariam dispostos a pagar para realizar a mudança é maior do que a quantidade de dinheiro que os perdedores teriam de ser compensados para não se opor para a mudança. (Os economistas são céticos sobre o que se aprende perguntando às pessoas quanto estariam dispostas a pagar e, em vez disso, tentam inferir quanto as pessoas estão dispostas a pagar indiretamente a partir dos fenômenos de mercado.) Quando **S** é uma melhoria potencial de Pareto em relação a **R**, diz-se que há um “benefício líquido” para a política de obtenção de **S**. De acordo com a análise de custo-benefício, entre as políticas elegíveis (que satisfaçam as restrições legais e morais), deve-se, em igualdade de condições, empregar aquela com o maior benefício líquido. Observe que a compensação é totalmente hipotética. As melhorias potenciais de Pareto resultam em vencedores e perdedores, cuja justiça ou injustiça é irrelevante para a análise de custo-benefício. Justiça ou beneficência podem exigir que a sociedade faça algo para mitigar desequilíbrios distributivos. Porque há um grande “bolo” de bens e serviços para satisfazer as preferências (uma vez que a compensação poderia ser paga e as preferências de todos mais bem satisfeitas), selecionar políticas com o maior benefício líquido serve à eficiência econômica (HICKS, 1939; KALDOR, 1939).

Apesar da importância prática da análise de custo-benefício, a técnica e a justificativa para ela esboçadas no parágrafo anterior são problemáticas. Uma dificuldade técnica é que é possível que **S** seja um aprimoramento potencial de Pareto sobre **R** e que **R** seja um aprimoramento potencial de Pareto sobre **S** (SCITOVSKY, 1941; SAMUELSON, 1950)! Isso significa que o fato de **S** ser uma melhoria potencial de Pareto sobre **R** não implica que haja um “bolo” econômico maior em **S** do que em **R**, porque não pode, é claro, haver um bolo econômico maior em **S** do que em **R** e um bolo econômico maior em **R** do que em **S**. Um segundo problema é que a disposição de pagar por alguma política e a quantia que alguém

exigiria em compensação se alguém se opusesse a ela dependem de quanta riqueza se possui, bem como de sua atitude em relação à política. A análise de custo-benefício pondera as preferências dos ricos mais do que as preferências dos pobres (BAKER, 1975). É possível compensar aproximadamente os efeitos de renda e riqueza (HARBURGER, 1978; FANKHAUSER *et al.*, 1997), mas é incômodo fazer isso, e a análise de custo-benefício é comumente empregada sem qualquer ajuste para riqueza ou renda.

Outra dificuldade séria para a economia do bem-estar tradicional, que tem estado como que escondida à vista de todos, é o fato de que as escolhas são indicadores imperfeitos de preferências, que por sua vez são indicadores imperfeitos do que melhora o bem-estar. Os mesmos fatos que mostram que a satisfação de preferências não constitui bem-estar (falsas crenças, falta de informação, preferências direcionadas a outros e não racionais) mostram que as escolhas e preferências às vezes são indicadores enganosos de bem-estar. Além disso, uma vez que se reconhece que as preferências são bons indicadores de bem-estar apenas se os agentes são bons julgadores do que os beneficiará, é necessário reconhecer que os agentes nem sempre são bons julgadores do que os beneficiará, mesmo quando têm todas as informações necessárias. Em alguns contextos, esses problemas podem ser menores. Por exemplo, as preferências das pessoas entre automóveis novos são amplamente autointeressadas, atenciosas e bem-informadas. Em outros contextos, como a proteção ambiental, as preferências por ignorar os problemas costumam ser mal-informadas, enquanto as preferências para agir normalmente não são autointeressadas. De qualquer forma, as preferências populares entre as políticas para tratar de problemas ambientais provavelmente não serão um bom guia para o bem-estar.

Ignorar esses problemas tem sido uma grande conveniência para a economia normativa. Se o que as pessoas escolhem revela suas preferências, o que por sua vez indica o que é bom para elas, então, como observado antes, a ação do governo para orientar as escolhas de alguém nunca pode tornar essa pessoa melhor e, portanto, não podem surgir dúvidas sobre endossar políticas paternalistas. Mas, seja aconselhável ou não, o paternalismo bem-sucedido não é impossível; e o trabalho recente de economistas comportamentais, que documentam uma ampla variedade de fraquezas deliberativas sistemáticas, colocou as questões relativas ao paternalismo de volta à mesa (ARIELY, 2009; KAHNEMAN, 2011). Alguns economistas têm procurado maneiras de identificar as “verdadeiras” preferências de um agente, conforme descrito por Infante *et al.* (2016). Outros argumentaram

que os formuladores de políticas devem respeitar as preferências dos agentes entre seus fins ou objetivos, ao mesmo tempo em que anulam as preferências entre os meios quando estes são distorcidos por mau julgamento ou falsas crenças (THALER; SUNSTEIN, 2008; LE GRAND; NEW, 2015). Além disso, a proposta de Thaler e Sunstein de que o governo explore métodos não coercitivos de influenciar as pessoas a fazerem melhores escolhas (“*nudges*” ou “empurrãozinhos”) tem sido popular entre os formuladores de políticas públicas e pode ter mudado a discussão filosófica do paternalismo do foco de Mill (1859) em evitar a coerção (SHIFFRIN, 2000; HAUSMAN; WELCH, 2010; LE GRAND; NEW, 2015).

6.3 Outras direções na economia normativa

Embora a economia do bem-estar e as preocupações com a eficiência dominem a economia normativa, eles não esgotam o assunto e, em colaboração com os filósofos, os economistas têm feito contribuições importantes para o trabalho contemporâneo em ética e filosofia social e política normativa. A Seção 5.2 e a Seção 5.3 deram algumas sugestões sobre as contribuições da teoria da escolha social e da teoria dos jogos. Além disso, economistas e filósofos trabalharam no problema de fornecer uma caracterização formal da liberdade de modo a trazer ferramentas de análise econômica para suportá-la (PATTANAIK; XU, 1990; SEN, 1988, 1990, 1991; CARTER, 1999; SUGDEN, 2018). Outros desenvolveram caracterizações formais de funções de bem-estar social que priorizam os interesses dos menos favorecidos ou que favorecem a igualdade de recursos, oportunidades e resultados e que separam a responsabilidade individual e social pelas desigualdades (PAZNER; SCHMEIDLER, 1974; VARIAN, 1974, 1975, ROEMER, 1986b, 1987; FLEURBAEY, 1995, 2008; FLEURBAEY; MANIQUET, 2014; GREAVES, 2015; MCCARTHY, 2015, 2017). John Roemer colocou a modelagem econômica contemporânea para trabalhar para oferecer caracterizações precisas da noção de **exploração** (1982). Amartya Sen e Martha Nussbaum desenvolveram não apenas novas interpretações das preocupações próprias da economia normativa em termos de **capabilidades** (SEN, 1992; NUSSBAUM; SEN, 1993; NUSSBAUM, 2000), que Sen vinculou a caracterizações de igualitarismo e a medidas operacionais de privação (SEN, 1999). Existem muitas interações atuais entre a economia normativa e a filosofia moral. Veja também as entradas sobre **Libertarianismo**, **Paternalismo**, **Igualitarismo** e **Justiça Econômica**, no site da SEP, em inglês.

7. Conclusões

As fronteiras entre a economia e a filosofia preocupadas com a metodologia, a racionalidade, a ética e a filosofia normativa social e política estão fervilhando de atividade. Esta atividade é diversa e preocupada com questões muito diferentes. Embora muitas delas estejam relacionadas, a filosofia da economia não é um empreendimento único e unificado. É uma coleção de investigações separadas, ligadas umas às outras por conexões entre as questões e pela influência dominante dos modelos e técnicas econômicas *mainstream*.

Referência Bibliográfica

A bibliografia a seguir não é abrangente. Ela normalmente evita citações separadas para ensaios metodológicos em coleções. Não lista separadamente os ensaios sobre metodologia econômica de edições especiais sobre filosofia e economia. Muitos ensaios sobre filosofia da economia podem ser encontrados nas revistas *Economics and Philosophy*, no *Journal of Economic Methodology* e na série anual *Research in the History of Economic Thought and Methodology* [Pesquisa sobre a História do Pensamento e da Metodologia Econômica].

Pode ser do interesse dos leitores consultar o *Journal of Economic Methodology*, vol. 8, n. 1, de março de 2001, Simpósio do Milênio sobre “The Past, Present and Future of Economic Methodology” [O Passado, o Presente e o Futuro da Metodologia Econômica] e Binder *et al.*, 2016. Para uma visão geral enciclopédica da metodologia econômica, consulte o *Handbook of Economic Methodology* [Manual de Metodologia Econômica], editado por Davis, Hands e Mäki. Para obter uma bibliografia abrangente de trabalhos sobre metodologia econômica até 1988, consulte Redman (1989). Ensaios de revistas de economia são indexados no *The Journal of Economic Literature* [Revista de Literatura Econômica], e o *Index of Economic Articles in Journal and Collective Volumes* [Índice de Artigos Econômicos em Revistas e Coleções] também indexa coleções. Desde 1991, os trabalhos sobre metodologia podem ser encontrados sob o número B4. Trabalhos sobre ética e economia podem ser encontrados sob os números A13, D6 e I3. As discussões sobre a racionalidade e a teoria dos jogos podem ser encontradas em A1, C7, D00, D7, D8 e D9.

Metodologia Econômica

- ACKROYD, S.; FLEETWOOD, S. **Critical Realist Applications in Organisation and Management Studies**. London: Routledge, 2004.
- ALCHIAN, A. Uncertainty, Evolution and Economic Theory. **Journal of Political Economy**, v. 57, p. 211-21, 1950.
- ALT, J.; LEVI, M.; OSTROM, E. (ed.). **Competition and Cooperation: Conversations with Nobelists about Economics and Political Science**. New York: Russell Sage Foundation, 1999.
- AMARIGLIO, J.; CULLENBERG, S.; RUCCIO, D. (ed.). **Post-Modernism, Economics and Knowledge**. London: Routledge, 2001.
- ANGYROUS, G. Refutation or Comparison? **British Journal for the Philosophy of Science**, v. 17, p. 279-296, 1967.
- ARDALAN, K. **Paradigms in Political Economy**. London: Routledge, 2016.
- ARIELY, D. **Predictably Irrational: The Hidden Forces That Shape Our Decisions**. New York, Harper, 2009.
- ARROW, K. Economic Equilibrium. **International Encyclopedia of the Social Sciences**. New York: Macmillan, 1968, p. 376-389.
- ARROW, K. **The Limits of Organization**. New York: Norton, 1974.
- ARROW, K.; HAHN, F. **General Competitive Analysis**. San Francisco: Holden-Day, 1971.
- ARUKA, Y. **Evolutionary Foundations of Economic Science: How Can Scientists Study Evolving Economic Doctrines from the Last Centuries?** New York: Springer, 2015.
- AUERBACH, P. **Socialist Optimism: An Alternative Political Economy for the Twenty-First Century**. New York: Palgrave Macmillan, 2016.
- AYER, A. **Language, Truth and Logic**. 2. ed. repr. New York: Dover, 1946.
- BACKHOUSE, R. (ed.). **New Perspectives on Economic Methodology**. London: Routledge, 1994.
- BACKHOUSE, R. **Truth and Progress in Economic Knowledge**. Cheltenham: Edward Elgar, 1997.
- BACKHOUSE, R. **The Rise and fall of Popper and Lakatos in Economics**, 2012. DOI:
- BACKHOUSE, R. **The Puzzle of Modern Economics**. Cambridge: Cambridge University Press, 2010.
- BACKHOUSE, R.; HAUSMAN, D.; MÄKI, U. (ed.). **Economics and Methodology: Crossing Boundaries**. London: Palgrave Macmillan, 1998.

- BACKHOUSE, R.; MEDEMA, S. The Definition of Economics. **Journal of Economic Perspectives**, v. 23, n. 1, p. 221-233, 2009.
- BACKHOUSE, R.; FONTAINE, P. (ed.). **The Unsocial Social Science? Economics and Neighboring Disciplines since 1945**. Durham: Duke University Press, 2010.
- BACKHOUSE, R. E.; BATEMAN, B. W. **Capitalist Revolutionary**: John Maynard Keynes. London: Harvard University Press, 2011.
- BALZER, W.; HAMMINGA, B. (ed.). **Philosophy of Economics**. Dordrecht: Kluwer-Nijhoff, 1989.
- BARDSLEY, N.; CUBITT, R. **Experimental Economics**: Rethinking the Rules. Princeton: University Press, 2009.
- BARKER, D.; KUIPER, E. (ed.). **Toward a Feminist Philosophy of Economics**. London: Routledge, 2003.
- BATEMAN, B. Keynes, Induction and Econometrics. **History of Political Economy**, v. 22, p. 359-379, 1990.
- BAUMBERGER, J. No Kuhnian Revolutions in Economics. **Journal of Economic Issues**, v. 11, p. 1-20, 1977.
- BAVETTA, S.; NAVARRA, P.; MAIMONE, D. **Freedom and the Pursuit of Happiness**: An Economic and Political Perspective. Cambridge: Cambridge University Press, 2014.
- BEAR, D.; ORR, D. Logic and Expediency in Economic Theorizing. **Journal of Political Economy**, v. 75, p. 188-196, 1967.
- BECKER, G. Irrational Behavior and Economic Theory. **Journal of Political Economy**, v. 70, p. 1-13, 1962.
- BECKER, G. **The Economic Approach to Human Behavior**. Chicago: University of Chicago Press, 1976.
- BECKERT, J. **Imagined Futures**: Fictional Expectations and Capitalist Dynamics. Cambridge: Harvard University Press, 2016.
- BEED, C. Philosophy of Science and Contemporary Economics. **Journal of Post Keynesian Economics**, v. 13, p. 459-494, 1991.
- BEGG, D. **The Rational Expectations Revolution in Macroeconomics**: Theories and Evidence. Baltimore: Johns-Hopkins University Press, 1982.
- BELL, D.; KRISTOL, I. (ed.). **The Crisis in Economic Theory**. New York: Basic Books, 1981.
- BEN-NER, A.; PUTTERMAN, L. (ed.). **Economics, Values and Organization**. Cambridge: Cambridge University Press, 1998.

- BERGER, L. Economics and Hermeneutics. **Economics and Philosophy**, v. 5, p. 209-234, 1989.
- BHASKAR, R.; ARCHER, M.; COLLIER, A.; LAWSON, T.; NORRIE, A. (ed.). **Critical Realism**. London: Routledge, 1998.
- BILLINGS, D. **Introducing Economic Actualism**. 2. ed. Bloomington: Author House, 2016.
- BINDER, C.; HEILMANN, C.; VROMEN, J. **The Future of the Philosophy of Economics**. London: Routledge, 2016.
- BIRKS, S. **Rethinking Economics: From Analogies to the Real World**. New York: Springer, 2015.
- BIRNER, J. **Strategies and Programmes in Capital Theory: A Contribution to the Methodology of Theory Development**. Dissertation, University of Amsterdam, 1990.
- BLAUG, M. **The Cambridge Revolution. Success or Failure?** London: Institute of Economic Affairs, 1975.
- BLAUG, M. Kuhn versus Lakatos or, Paradigms versus Research Programmes in the History of Economics. In: LATSIS, S. (ed.). **Method and Appraisal in Economics**. Cambridge: Cambridge University Press, 1976, p. 149-180.
- BLAUG, M. **The Methodology of Economics: Or How Economists Explain**. Cambridge: Cambridge University Press, 1980a.
- BLAUG, M. **A Methodological Appraisal of Marxian Economics**. Amsterdam: North-Holland, 1980b.
- BLINDER, A. The Economics of Brushing Teeth. **Journal of Political Economy**, v. 82, p. 887-891, 1974.
- BLINDER, A. Learning by Asking Those Who are Doing. **Eastern Economic Journal**, v. 16, 297-306, 1990.
- BOEHM, S. *et. al.* (ed.). **Is There Progress in Economics? Knowledge, Truth and the History of Economic Thought**. Cheltenham: Edward Elgar, 2002.
- BOETTKE, P. J. (ed.). **Handbook on Contemporary Austrian Economics**. Cheltenham: Elgar, 2010.
- BOETTKE, P.; COYNE, C. (ed.). **The Oxford Handbook of Austrian Economics**. New York: Oxford University Press, 2015.
- BOLAND, L. A Critique of Friedman's Critics. **Journal of Economic Literature**, v. 17, p. 503-522, 1979.
- BOLAND, L. **The Foundations of Economic Method**. London: George Allen & Unwin, 1982.

- BOLAND, L. **The Methodology of Economic Model Building**: Methodology after Samuelson. London: Routledge, 1989.
- BOLAND, L. **The Principles of Economics**: Some Lies my Teachers Told Me. London: Routledge, 1992.
- BOLAND, L. **Critical Economic Methodology**: A Personal Odyssey. London: Routledge, 1997.
- BOLAND, L. **Model Building in Economics**: Its Purposes and Limitations. New York: Cambridge University Press, 2014.
- BOLAND, L. **Equilibrium Models in Economics**: Purposes and Critical Limitations. New York: Oxford University Press, 2017.
- BONAR, J. **Philosophy and Political Economy**. London: Allen & Unwin, 1967.
- BOULDING, K. **Economics as a Science**. New York: McGraw-Hill, 1970.
- BOULIER, B. Pisces Economicus: The Fish as Economic Man. **Economics and Philosophy**, v. 7, p. 83-86, 1991.
- BOUMANS, M. **Science Outside the Laboratory. Measurement in Field Science and Economics**. Oxford: Oxford University Press, 2015.
- BOUMANS, M.; DAVIS, J. B. **Economic Methodology**: Understanding Economics as a Science. 2. ed. New York: Palgrave Macmillan, 2015.
- BOYLAN, T.; O'GORMAN, P. **Beyond Rhetoric and Realism in Economics**: Towards a Reformulation of Economic Methodology. London: Routledge, 1995.
- BOYLAN, T.; O'GORMAN, P. (ed.). **Popper and Economic Methodology**: Contemporary Challenges. London: Routledge, 2007.
- BRANAS-GARZA, P.; CABRALES, A. (ed.). **Experimental Economics**. 2 v. London: Macmillan, 2015.
- BRAY, J. The Logic of Scientific Method in Economics. **Journal of Economic Studies**, v4: 1-28. 1977
- BRENNAN J.; JAWORSKI, P. **Markets Without Limits**: Moral Virtues and Commercial Interests. London: Routledge, 2016.
- BRENNAN, T. A Methodological Assessment of Multiple Utility Frameworks. **Economics and Philosophy**, v. 5, p. 189-208, 1989.
- BRONFENBRENNER, M. The Structure of Revolutions in Economic Thought. **History of Political Economy**, v. 3, p. 136-151, 1971.
- BROOME, J. **Climate Matters**: Ethics in a Warming World. New York: Norton, 2012.
- BROUSSEAU, E.; GLACHANT, J. (ed.). **New Institutional Economics**: A Guidebook. Cambridge: Cambridge University Press, 2008.

- BROWN, A.; FLEETWOOD, S. **Critical Realism and Marxism**. London: Routledge, 2003.
- BRUNI, L. **The Genesis and Ethos of the Market**. New York: Palgrave-Macmillan, 2012.
- BRUNNER, K. Assumptions and the Cognitive Quality of Theories. **Synthese**, v. 20, 501–25. 1969
- BRZEZINSKI, J.; CONIGLIONE, F.; KUIPERS, R.; NOWAK, L. (eds.). **Idealization I: General Problems** (Poznan Studies in the Philosophy of the Sciences and Humanities, v. 16). Amsterdam: Rodopi, 1990.
- BUCHANAN, J. *Ceteris Paribus*: Some Notes on Methodology. **Southern Economic Journal**, v. 24, p. 259-270, 1958.
- BUCHANAN, J. **The Limits of Liberty**: Between Anarchy and the Leviathan. Chicago: University of Chicago Press, 1975.
- BUCHANAN, J. **What Should Economists Do?** Indianapolis: Liberty Press, 1979.
- BUCHANAN, J. **Economics from the Outside In Better than Plowing and Beyond**. College Station: Texas A&M University Press, 2007.
- BUCHANAN, J.; VANBERG, V. The Market as a Creative Process. **Economics and Philosophy**, v. 7, 167-186, 1979.
- BUECHNER, M. N. **Objective Economics**: How Ayn Rand's Philosophy Changes Everything about Economics. Lanham: Rowman and Littlefield, 2011.
- CAIRNES, J. **The Character and Logical Method of Political Economy**. 2. ed. repr. New York: A. M. Kelley, 1965.
- CALDWELL, B. **Beyond Positivism**: Economic Methodology in the Twentieth Century. London: Allen & Unwin, 1982.
- CALDWELL, B. The Neoclassical Maximization Hypothesis: Comment. **American Economic Review**, v. 75, p. 824-827, 1983.
- CALDWELL, B. Does Methodology Matter? How Should It Be Practiced? **Finnish Economic Papers**, v. 3, p. 64-71, 1990.
- CALDWELL, B. Clarifying Popper. **Journal of Economic Literature**, v. 29, p. 1-33, 1991.
- CALDWELL, B. (ed.). **Appraisal and Criticism in Economics**. London: Allen & Unwin, 1984.
- CALDWELL, B. (ed.). **The Philosophy and Methodology of Economics**. Cheltenham: Edward Elgar, 1993.
- CAMERER, C. **Behavioral Game Theory**: Experiments in Strategic Interaction. Princeton: Princeton University Press, 2003.
- CAMERER, C. Neuroeconomics: Using Neuroscience to Make Economic Predictions. **Economic Journal**, v. 117, C26-42, 2007.

- CAMERER, C.; LOEWENSTEIN, G.; RABIN, M. (eds.). **Advances in Behavioral Economics**. Princeton: Princeton University Press, 2003.
- CAMERER, C.; LOEWENSTEIN, G.; PRELEC, D. Neuroeconomics: How Neuroscience Can Inform Economics. **Journal of Economic Literature**, v. 43, p. 9-64, 2005.
- CAMERER, C.; COHEN, J.; FEHR, E.; GLIMCHER, P.; LAIBSON, D.; LOEWENSTEIN, G.; MONTAGUE, R.; Neuroeconomics. *In*: KAGEL, J.; ROTH, A. (ed.). **The Handbook of Experimental Economics**. 2. ed. Princeton: Princeton University Press, 2008.
- CAMERER, C.; LOEWENSTEIN, (ed.). **Advances in Behavioral Economics**. Princeton: Princeton University Press. 2013
- CAPLIN, A.; SCHOTTER, A. (ed.). **Handbook of Economic Methodology**. New York: Oxford University Press, 2008.
- CARPENTER, J.; HARRISON, G.; LIST, J. (ed.). **Research in Experimental Economics**. Bingley: Emerald Publishing, 2005.
- CARTER, M.; MADDOCK, R. **Rational Expectations**: Macroeconomics for the 1980s? London: Macmillan, 1984.
- CHIPMAN, J. The Nature and Meaning of Equilibrium in Economic Theory. *In*: MARTINDALE, D. (ed.). **Functionalism in the Social Sciences**: The Strength and Limits of Functionalism in Anthropology, Economics, Political Science and Sociology. Philadelphia: American Academy of Political and Social Science, 1965, p. 35-64.
- CLARK, A.; FLÈCHE, S. **The Origins of Happiness**: The Science of Well-Being over the Life Course. Princeton: Princeton University Press, 2018.
- CLIFT, E. M. (ed.). **How Language Is Used to Do Business**: Essays on the Rhetoric of Economics. Lewiston: Edwin Mellen Press, 2008.
- COASE, R. The Problem of Social Cost. **Journal of Law and Economics**, v. 3, p. 1-30, 1960.
- COATS, A. Is There a 'Structure of Scientific Revolutions' in Economics? **Kyklos**, v. 22, p. 289-294, 1969.
- COATS, A. The Methodology of Economics: Some Recent Contributions. **Kylos**, v. 35, p. 310-321, 1982.
- CODDINGTON, A. Positive Economics. **Canadian Journal of Economics**, v. 5, p1-15. 1972.
- CODDINGTON, A. The Rationale of General Equilibrium Theory. **Economic Inquiry**, v. 13, p. 539-558, 1975.

- COHEN, I.; DANIELS, N.; EYAL, N. (ed.). **Identified Versus Statistical Lives: An Interdisciplinary Perspective**. New York: Oxford University Press, 2015.
- COLANDER, D.; KLAMER, A. The Making of An Economist. **Journal of Economic Perspectives**, v. 1, p. 95-112, 1987.
- COLE, K.; CAMERON, J.; EDWARDS, C. **What Economists Disagree**. 2. ed. London: Longmans, 1991.
- COLEMAN, J. Economics and the Law: A Critical Review of the Foundations of the Economic Approach to Law. **Ethics**, v. 94, p. 649-679, 1984.
- COLLINS, H. The Meaning of Replication and the Science of Economics. **History of Political Economy**, v. 23, p. 123-142, 1991.
- COOK, S. J. **The Intellectual Foundation of Alfred Marshall's Economic Science: A Rounded Globe of Knowledge**. Cambridge: Cambridge University Press, 2009.
- COOTER, R.; RAPPOPORT, P. Were the Ordinalists Wrong About Welfare Economics? **Journal of Economic Literature**, v. 22, p. 507-530, 1984.
- COYNE, C.; STORR, V. (ed.). **New Thinking in Austrian Political Economy**. Bingley: Emerald Publishing, 2015.
- CRESPO, R. **Philosophy of the Economy: An Aristotelian Approach**, New York: Springer. 2013
- CROSS, R. The Duhem-Quine Thesis, Lakatos and the Appraisal of Theories in Macroeconomics. **Economic Journal**, v. 92, p. 320-340, 1982.
- CULLENBERG, S.; AMARIGLIO, J.; RUCCIO, F. (ed.). **Postmodernism, Economics, and Knowledge**. London: Routledge, 2001.
- CYERT, R.; GRUNBERG, E. Assumption, Prediction and Explanation in Economics. *In*: CYERT, R. M.; MARCH, J. G. **A behavioral theory of the firm**. Englewood Cliffs: Prentice-Hall, 1963, p. 298-311.
- CYERT, R.; POTTINGER, G. Towards a Better Micro-economic Theory. **Philosophy of Science**, v. 46, p. 204-222, 1979.
- DANNER, P. **The Economic Person**. Lanham: Rowman and Littlefield, 2002.
- D'AUTUME, A.; CARTELIER, J. (ed.). **Is Economics Becoming a Hard Science?** Cheltenham: Edward Elgar, 1997.
- DAVIES, W. **The Happiness Industry: How the Government and Big Business Sold us Well-Being**, London: Verso, 2015.
- DAVIS, J. **The Theory of the Individual in Economics: Identity and Value**. London: Routledge, 2003.
- DAVIS, J. **Individuals and Identity in Economics**. Cambridge: Cambridge University Press, 2011.

- DAVIS, J. (ed.). **Recent Developments in Economic Methodology**. Cheltenham: Edward Elgar, 2006.
- DAVIS, J.; BOUMANS, M. **Economic Methodology: Understanding Economics as a Science**. London: Palgrave, 2010.
- DAVIS, J.; WADE HANDS, D. (ed.). **The Elgar Companion to Recent Economic Methodology**. Cheltenham: Edward Elgar, 2011.
- DAVIS, J.; WADE HANDS, D. (ed.). **Reflexivity and Economics: George Soros's theory of reflexivity and the methodology of economic science**. London: Routledge, 2015.
- DAVIS, J.; WADE HANDS, D.; MÄKI, U. (eds.). **The Handbook of Economic Methodology**. Cheltenham: Edward Elgar, 1998.
- DAVIS, J.; MARCIANO, A.; RUNDE, J. (ed.). **The Elgar Companion to Economic and Philosophy**. Cheltenham: Edward Elgar, 2004.
- De ALESSI, L. Reversals of Assumptions and Implications. **Journal of Political Economy**, v. 79, p. 867-877, 1971.
- DEATON, A. Instruments, Randomization, and Learning about Development. **Journal of Economic Literature**, v. 48, p. 424-455, 2010.
- DEBREU, G. **Theory of Value**. New York: Wiley, 1959.
- DEKKER, E. **The Viennese Students of Civilization: The Meaning and Context of Austrian Economics Reconsidered**. New York: Cambridge University Press, 2016.
- DELORME, R. **Deep Complexity and the Social Sciences: Experience, Modelling and Operationality (New Horizons in Institutional and Evolutionary Economics)**. Northampton: Elgar, 2010.
- DE MARCHI, N. The Empirical Content and Longevity of Ricardian Economics. **Economica**, v. 37, p. 257-276, 1970.
- DE MARCHI, N. (ed.). **The Popperian Legacy in Economics**. Cambridge: Cambridge University Press, 1988.
- DE MARCHI, N. (ed.). **Post-Popperian Methodology of Economics: Recovering Practice**. Boston: Kluwer, 1992.
- DE MARCHI, D.; BLAUG, M. (ed.). **Appraising Modern Economics: Studies in the Methodology of Scientific Research Programs**. Cheltenham: Edward Elgar, 1991.
- DEVROEY, M. The Base Camp Paradox: A Reflection the the Place of Tâtonnement in General Equilibrium Theory. **Economics and Philosophy**, v. 6, p. 235-254, 1990.

- DIESING, P. **Science and Ideology in the Policy Sciences**. New York: Aldine, 1982.
- DIETRICH, F.; LIST, C. Mentalism versus Behaviourism in Economics: A Philosophy-of-Science Perspective. **Economics and Philosophy**, v. 32, p. 249-282, 2016.
- DIETSCH, P. **Catching Capital: The Ethics of Tax Competition**. Oxford: Oxford University Press, 2015.
- DILLARD, D. Revolutions in Economic Theory. **Southern Economic Journal**, 44: 705–24. 1978.
- DOLAN, E. (ed.). **The Foundations of Modern Austrian Economics**. Kansas City: Sheed & Ward, 1976.
- DOLAN, P.; KAHNEMAN, D. **Happiness by Design: Change What You Do, Not How You Think**. New York: Avery Press, 2014.
- DOPFER, K.; POTTS, J. **The General Theory of Economic Evolution**. New York: Routledge, 2008.
- DOPFER, K.; POTTS, J. (ed.). **The New Evolutionary Economics**, 3 v. Cheltenham, U.K.: Elgar, 2014.
- DOW, S. **Macroeconomic Thought: A Methodology Approach**. Oxford: Blackwell, 1985.
- DOW, S. **Economic Methodology: An Inquiry**. Oxford: Oxford University Press, 2002.
- DOW, S. **Foundations for New Economic Thinking**. London: Palgrave Macmillan, 2012.
- DUFLO, E.; BANERJEE, A. **Poor Economics**. New York: Public Affairs, 2011.
- DUFLO, E.; BANERJEE, A. **Handbook of Field Experiments**. Amsterdam: North-Holland, 2017.
- DUGGER, W. Methodological Differences between Institutional and Neoclassical Economics. **Journal of Economic Issues**, v. 13, p. 899-909, 1979.
- DURLAUF, S.; BLUME, L. (ed.). **Behavioural and Experimental Economics**. London: Macmillan, 2009.
- DYKE, C. **Philosophy of Economics**. Englewood Cliffs: Prentice-Hall, 1981.
- EDWARDS, P.; MAHONEY, J.; VINCENT, S. (ed.). **Studying Organizations Using Critical Realism: A Practical Guide**. Oxford: Oxford University Press, 2014.
- EICHNER, A. Why Economics Is not yet a Science. In: EICHNER, A. (ed.). **Why Economics Is not yet a Science**. Armonk: M.E. Sharpe, 1983, p. 205-241.
- ELSTER, J. **Making Sense of Marx**. Cambridge: Cambridge University Press, 1985.
- ELSTER, J.; ROEMER, J. (ed.). **Interpersonal Comparisons of Well-Being**. Cambridge: Cambridge University Press, 1991.
- ENGLE, R.; HENDRY, D.; RICHARD, J. Exogeneity. **Econometrica**, v. 51, p. 277-304, 1983.

- ENGELSKIRCHEN, H. **Capital as a Social Kind: Definitions and Transformations in the Critique of Political Economy.** New York: Routledge, 2011.
- EPSTEIN, B. **The Ant Trap: Rebuilding the Foundations of the Social Sciences.** New York: Oxford University Press, 2015.
- ETZIONI, A. **The Moral Dimension. Toward a New Economics.** New York: Macmillan, 1988.
- ETZIONI, A. **Happiness is the Wrong Metric: A Liberal Communitarian Response to Populism.** New York: Springer, 2018.
- FAMA, E. Agency Problems and the Theory of the Firm. **Journal of Political Economy**, v. 88, p. 288-307, 1980.
- FERBER, M.; NELSON, J. (ed.). **Feminist Economics Today: Beyond Economic Man.** Chicago: University of Chicago Press, 2003.
- FISHER, R. **The Logic of Economic Discovery: Neoclassical Economics and the Marginal Revolution.** New York: New York University Press, 1986.
- FLEETWOOD, S. (ed.). **Critical Realism in Economics: Development and Debate.** London: Routledge, 1999.
- FOLDVARY, F. (ed.). **Beyond Neoclassical Economics: Heterodox Approaches to Economic Theory.** Cheltenham: Edward Elgar, 1996.
- FOX, G. **Reason and Reality in the Methodologies of Economics.** Cheltenham: Edward Elgar, 1997.
- FRANK, R. **Passions within Reason: The Strategic Role of the Emotions.** New York: W. W. Norton, 1988.
- FRANKFURTER, G.; MCGOUN, E. (ed.). **From Individualism to the Individual: Ideology and Inquiry in Financial Economics.** Aldershot: Ashgate, 2002.
- FRASER, L. **Economic Thought and Language. A Critique of Some Fundamental Concepts.** London: A & C Black, 1937.
- FREEDMAN, C. F. **Chicago Fundamentalism: Ideology and Methodology in Economics.** Hackensack: World Scientific, 2008.
- FREY, B. **Economics as a Science of Human Behaviour: Towards a New Social Science Paradigm.** 2. ed. Dordrecht: Kluwer, 1999.
- FREY, B. **Happiness: A Revolution in Economics.** Cambridge: MIT Press, 2010.
- FREY, B. **Economics of Happiness.** New York: Springer, 2018.
- FRÉCHETTE, G.; SCHOTTER, A. **Handbook of Experimental Economic Methodology.** Oxford: Oxford University Press, 2015.
- FRIEDMAN, M. **The Methodology of Positive Economics. Essays in Positive Economics.** Chicago: University of Chicago Press, 1953, p. 3-43.

- FRIEDMAN, M. Leon Walras and His Economic System. *In*: RIMA, I. (ed.). **Readings in the History of Economic Theory**. New York: Holt, Rinehart & Winston, 1970, p. 145-153.
- FULLBROOK, E. **A Guide to What's Wrong with Economics**. New York: Anthem Press, 2004.
- FULLBROOK, E. (ed.). **Intersubjectivity in Economics: Agents and Structures**. London: Routledge, 2001.
- FULLBROOK, E. (ed.). **The Crisis in Economics**. London: Routledge, 2003.
- FULLBROOK, E. (ed.). **Ontology and Economics: Tony Lawson and His Critics**. New York: Routledge, 2009.
- FURUBOTN, E. G.; RICHTER, R. (ed.). **The New Institutional Economics of Markets**. Cheltenham: Elgar, 2010.
- GANI, M. **Foundations of Economic Science**. Dhaka: Scholars, 2003.
- GAUS, G. **On Philosophy, Politics, and Economics**. Belmont: Wadsworth, 2008.
- GAY, D. **Reflexivity and Development Economics: Methodology, Policy and Practice**. New York: Palgrave Macmillan, 2009.
- GEORGE, D. (ed.). Issues in Heterodox Economics. **Journal of Economic Surveys**, v. 21, n. 3. Oxford: Blackwell, 2007.
- GERRARD, B. On Matters Methodological in Economics: Review Article. **Journal of Economic Surveys**, v. 4, p. 197-219, 1990.
- GERRARD, B. The Scientific Basis of Economics: A Review of the Methodological Debates in Economics and Econometrics. **Scottish Journal of Political Economy**, v. 42, p. 201-220, 1995.
- GEORGESCU-ROEGEN, N. Methods in Economic Science. **Journal of Economic Issues**, v. 13, p. 317-328, 1979.
- GEWEKE, J. Causality, Exogeneity and Inference. *In*: HILDENBRAND, W. (ed.). **Advances in Econometrics**. Cambridge: Cambridge University Press, 1982.
- GIBBARD, A.; VARIAN, H. Economic Models. **Journal of Philosophy**, v. 75, p. 664-677, 1978.
- GIOCOLI, N. **Modeling Rational Agents: From Interwar Economics to Early Modern Game Theory**. Cheltenham: Edward Elgar, 2003.
- GLIMCHER, P. **Foundations of Neuroeconomic Analysis**. New York: Oxford University Press, 2010.
- GLIMCHER, P.; CAMERER, C.; POLDRACK, R.; FEHR, E. (ed.). **Neuroeconomics: Decision Making and the Brain**. Amsterdam: Elsevier, 2009.

- GLIMCHER, P.; FEHR, E. (ed.). **Neuroeconomics**. 2. ed. New York: Academic Press, 2013.
- GONZALEZ, W. **Scientific Prediction and Economics: A Philosophical Analysis**. Pittsburgh: University of Pittsburgh Press, 2008.
- GORDON, D. Operational Propositions in Economic Theory. **Journal of Political Economy**, v. 63, p. 150-161, 1955.
- GRANGER, C. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. **Econometrica**, v. 37, p. 424-438, 1969.
- GRANGER, C. Testing for Causality: A Personal Viewpoint. **Journal of Economic Dynamics and Control**, v. 2, p. 329-352, 1980.
- GRANOVETTER, M. Economics and Social Structure: The Problem of Embeddedness. **American Journal of Sociology**, v. 91, p. 481-510, 1985.
- GRAPARD, U.; HEWITSON, G. (ed.). **Robinson Crusoe's Economic Man: A Construction and Deconstruction**. New York: Routledge, 2011.
- GREEN, E. On the Role of Fundamental Theory in Positive Economics. In: PITT, J. (ed.). **Philosophy in Economics**. Dordrecht: Reidel, 1981, p. 5-15.
- GRÖNKVIST, U. **Economic Methodology: Patterns of Reasoning and the Structure of Theories**. Lund: University of Lund, 1992.
- GROSSBARD-SHECHTMAN, S.; CLAGUE, C. (ed.). **The Expansion of Economics: Toward a more Inclusive Social Science**. Armonk: M. E. Sharpe, 2002.
- GRUNDBERG, E. 'Complexity' and 'Open Systems' in Economic Discourse. **Journal of Economic Issues**, v. 12, p. 541-560, 1978.
- GRÜNE-YANOFF, T. Learning from Minimal Economic Models. **Erkenntnis**, v. 70, p. 81-99, 2009.
- GUALA, F. Artefacts in Experimental Economics: Preference Reversals and the Becker-DeGroot-Marschak Mechanism. **Economics and Philosophy**, v. 16, p. 47-75, 2000a.
- GUALA, F. The Logic of Normative Falsification: Rationality and Experiments in Decision Theory. **Journal of Economic Methodology**, v. 7, p. 59-93, 2000b.
- GUALA, F. **The Methodology of Experimental Economics**. Cambridge: Cambridge University Press, 2005.
- GUALA, F. **Understanding Institutions: The Science and Philosophy of Living Together**. Princeton: Princeton University Press, 2016.
- GUGERTY, M. K.; KARLAN, D. **The Goldilocks Challenge: Right-Fit Evidence for the Social Sector**, Oxford: Oxford University Press, 2018.

- GUL, F.; PESANDORFER, W. The Case for Mindless Economics. *In*: CAPLIN, A.; SCHOTTER, A. (ed.). **Handbook of Economic Methodology**. New York: Oxford University Press, 2008, p. 3-39.
- GURAK, H. **Economic Growth and Development 2: Complementary Articles in the Pursuit of Economic Realities**. Frankfurt am Main: Peter Lang, 2018.
- GUSTAFSSON, B.; KNUDSEN, C.; MÄKI, U. (ed.). **Rationality, Institutions and Economic Methodology**. London: Routledge, 1993.
- HAAVELMO, T. The Probability Approach in Econometrics. **Econometrica**, v. 12, supp. 1-118, 1944.
- HAHN, F. The Winter of Our Discontent. **Economica**, v. 40, p. 322-330, 1973.
- HAGEDORN, H. **A Model of Austrian Economics**. Wiesbaden: Springer Gabler, 2015.
- HAHN, F.; HOLLIS, (ed.). **Philosophy and Economic Theory**. Oxford: Oxford University Press, 1979.
- HALL, R.; HITCH, C. Price Theory and Business Behaviour. **Oxford Economic Papers**, v. 2, p. 12-45, 1939.
- HAMMINGA, B. **Neoclassical Theory Structure and Theory Development: An Empirical-Philosophical Case Study Concerning the Theory of International Trade**. Boston: Springer, 1983.
- HAMMINGA, B.; DEMARCHI, N. (ed.). **Idealization in Economics**. Amsterdam: Rodopi, 1994.
- HAMMOND, J. D. Frank Knight's Antipositivism. **History of Political Economy**, v. 23, p. 355-381, 1991.
- HAMMOND, J. D. An Interview with Milton Friedman on Methodology. *In*: SAMUELS (ed.). **Research in the history of economic thought and methodology : a research annual**. **Emerald**, v. 10, 1992, p. 91-118.
- HÄNDLER, E. The Logical Structure of Modern Neoclassical Static Microeconomic Equilibrium Theory. **Erkenntnis**, v. 15, p. 33-53, 1980.
- HANDS, D. W. Karl Popper and Economic Methodology. **Economics and Philosophy**, v. 1, p. 83-100, 1985a.
- HANDS, D. W. Second Thoughts on Lakatos. **History of Political Economy**, v. 17, p. 1-16, 1985b.
- HANDS, D. W. The Structuralist View of Economic Theories: The Case of General Equilibrium in Particular. **Economics and Philosophy**, v. 1, p. 303-336, 1985c.
- HANDS, D. W.. "Ad Hocness in Economics and the Popperian Tradition. *In*: Marchi, N. (ed.). **The Popperian Legacy in Economics and Beyond**. Cambridge: Cambridge University Press, pp. 121-137, 1988.

- HANDS, D. W. **Testing, Rationality and Progress**. Totowa: Rowman and Littlefield, 1992.
- HANDS, D. W. Social Epistemology Meets the Invisible Hand: Kitcher on the Advancement of Science. **Dialogue**, v. 34, p. 605-621, 1995.
- HANDS, D. W. **Reflection Without Rules: Economic Methodology and Contemporary Science Theory**, Cambridge: Cambridge University Press, 2001.
- HANDS, D. W.; MIROWSKI, P. Harold Hotelling and the Neoclassical Dream. In: BACKHOUSE, R.; HAUSMAN, D.; MÄKI, U. (ed.). **Economics and Methodology: Crossing Boundaries**. London: Palgrave Macmillan, 1998, p. 322-397.
- HARCOURT, G.; KRIESLER, P. (ed.). **The Oxford Handbook of Post-Keynesian Economics**. Oxford: Oxford University Press, 2013.
- HARDT, L. **Economics without Laws: Toward a New Philosophy of Economics**. Switzerland: Palgrave Macmillan, 2017.
- HARROD, R. Scope and Method of Economics. **Economic Journal**, v. 48, p. 383-412, 1938.
- HAUSMAN, D. **Capital, Profits, and Prices: An Essay in the Philosophy of Economics**. New York: Columbia University Press, 1981.
- HAUSMAN, D. Are There Causal Relations Among Dependent Variables? **Philosophy of Science**, v. 50, p. 58-81, 1983.
- HAUSMAN, D. Supply and Demand Explanations and their *Ceteris Paribus*, Clauses. **Review of Political Economy**, v. 2, p. 168-186, 1990.
- HAUSMAN, D. **Essays on Philosophy and Economic Methodology**. Cambridge: Cambridge University Press, 1992a.
- HAUSMAN, D. **The Inexact and Separate Science of Economics**. Cambridge: Cambridge University Press, 1992b.
- HAUSMAN, D. Problems with Realism in Economics. **Economics and Philosophy**, v. 14, p. 185-213, 1998.
- HAUSMAN, D. Why Look Under the Hood, 2008a. In: HAUSMAN, D. (ed.). **The Philosophy of Economics: An Anthology**. 3. ed. Cambridge: Cambridge University Press, 2008b, p. 217-221.
- HAUSMAN, D. Mistakes about Preferences in the Social Sciences. **Philosophy of the Social Sciences**, v. 41, p. 3-25, 2011.
- HAUSMAN, D. **Preference, Value, Choice, and Welfare**. New York: Cambridge University Press, 2012.
- HAUSMAN, D. Paradox Postponed. **Journal of Economic Methodology**, v. 20, p. 250-254, 2013.

- HAUSMAN, D. Much Ado about Models, Review of Mary Morgan, *The World in the Model: How Economists Work and Think*. **Journal of Economic Methodology**, v. 22, p. 241-246, 2015a.
- HAUSMAN, D. **Valuing Health: Well-Being, Freedom, and Suffering**. New York: Oxford University Press, 2015b.
- HAUSMAN, D. (ed.). **The Philosophy of Economics: An Anthology**. 3. ed. Cambridge: Cambridge University Press, 2008b.
- HAUSMAN, J. Fiscal Policy and Economic Recovery: The Case of the 1936 Veterans' Bonus. **American Economic Review**, v. 106, p. 1100-1143, 2016.
- HAYEK, F. Economics and Knowledge. **Economica**, v. 4, p. 33-54, 1937.
- HEILBRONER, R. On the Limited Relevance of Economics. **Public Interest**, v. 21, p. 80-93, 1970.
- HELM, D. Predictions and Causes: A Comparison of Friedman and Hicks on Method. **Oxford Economic Papers**, v. 36, supp., p. 118-134, 1984.
- HENDERSON, W.; DUDLEY-EVANS, T.; BACKHOUSE, R. (ed.). **Economics and Language**. London: Routledge, 1993.
- HENDRY, D. **Econometrics - Alchemy or Science?** Oxford: Blackwell, 1993.
- HERRMANN-PILLATH, C. **Foundations of Economic Evolution: A Treatise on the Natural Philosophy of Economics**. Cheltenham: Elgar, 2013.
- HERZOG, L. **Inventing the Market: Smith, Hegel, and Political Theory**. Oxford: Oxford University Press, 2013.
- HEY, J. D. **Experiments in Economics**. Oxford: Blackwell, 1991.
- HICKS, J. The Foundations of Welfare Economics. **Economic Journal**, v. 49, p. 696-712, 1939.
- HICKS, J. **Causality in Economics**. New York: Basic Books, 1979.
- HICKS, J.; ALLEN, R. A Reconsideration of the Theory of Value. **Economica**, v. 1, p. 52-76; p. 196-219, 1934.
- HIRSCH, A.; MARCHI, N. de. **Milton Friedman: Economics in Theory and Practice**. Ann Arbor: University of Michigan Press, 1990.
- HIRSCH, F. **The Social Limits to Growth**. Cambridge: Harvard University Press, 1976.
- HIRSCHMAN, A. Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse. **Economics and Philosophy**, v. 1, p. 7-22, 1985.
- HODGSON, B. **Economics as Moral Science**. Heidelberg: Springer, 2001.
- HODGSON, G. What Is the Essence of Institutional Economics? **Journal of Economic Issues**, v. 34, p. 317-329, 2000.

- HODGSON, G. **The Evolution of Institutional Economics**. London: Routledge, 2004.
- HODGSON, G. **From Pleasure Machines to Moral Communities: An Evolutionary Economics without Homo economicus**. London: University of Chicago Press, 2013.
- HODGSON, G. **Conceptualizing Capitalism: Institutions, Evolution, Future**. London: University of Chicago Press, 2016.
- HODGSON, G. M.; KNUDSEN, T. **Darwin's Conjecture: The Search for General Principles of Social and Economic Evolution**. Chicago: University of Chicago Press, 2010.
- HOLCOMBE, R. **Advanced Introduction to the Austrian School of Economics**. Northampton: Elgar, 2014.
- HOLLAND, J.; HOLYOAK, K.; NISBETT, R.; THAGARD, P. **Induction: Processes of Inference, Learning, and Discovery**. Cambridge: MIT Press, 1986.
- HOLLIS, M.; NELL, E. **Rational Economic Man: A Philosophical Critique of Neo-Classical Economics**. London: Cambridge University Press, 1975.
- HOOVER, K. **The New Classical Macroeconomics: A Sceptical Inquiry**. Oxford: Basil Blackwell, 1988.
- HOOVER, K. Econometrics as Observation: The Lucas Critique and the Nature of Econometric Inference. **Journal of Economic Methodology**, v. 1, p. 65-80, 1994.
- HOOVER, K. **Causality in Macroeconomics**. Cambridge: Cambridge University Press, 2001a.
- HOOVER, K. **The Methodology of Empirical Macroeconomics**. Cambridge: Cambridge University Press, 2001b.
- HORN, K. I. **Roads to Wisdom, Conversations with Ten Nobel Laureates in Economics**. Northampton: Edward Elgar, 2009.
- HORWITZ, S. **Hayek's Modern Family: Classical Liberalism and the Evolution of Social Institutions**. New York: St. Martin's Press, 2015.
- HULL, D. **Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science**. Chicago: University of Chicago Press, 1988.
- HUME, D. Of Money, Of the Balance of Trade. In: ROTWEIN, E. (ed.). **David Hume Writings on Economics**. Madison: University of Wisconsin Press, 1970.
- HUMPHRIES, J. (ed.). **Gender and Economics**. Aldershot: Edward Elgar, 1995.
- HUTCHISON, T. **The Significance and Basic Postulates of Economic Theory**. New York: A.M. Kelley, 1960.

- HUTCHISON, T. The Significance and Basic Postulates of Economic Theory: A Reply to Professor Knight. **Journal of Political Economy**, v. 49, p. 732-750, 1941.
- HUTCHISON, T. Professor Machlup on Verification in Economics. **Southern Economic Journal**, v. 22, p. 476-483, 1956.
- HUTCHISON, T. Methodological Prescriptions in Economics: A Reply. **Economica**, v. 27, p. 158-160, 1960.
- HUTCHISON, T. **Knowledge and Ignorance in Economics**. Chicago: University of Chicago Press, 1977.
- HUTCHISON, T. **On Revolutions and Progress in Economic Knowledge**. Cambridge: Cambridge University Press, 1978.
- HUTCHISON, T. **The Politics and Philosophy of Economics**: Marxians, Keynesians and Austrians, Oxford: Basil Blackwell, 1981.
- HUTCHISON, T. **On the Methodology of Economics and the Formalist Revolution**. Cheltenham: Edward Elgar, 2000.
- INFANTE, G.; LECOUEUX, G.; SUGDEN, R. Preference Purification and the Inner Rational Agent: A Critique of the Conventional Wisdom of Behavioural Welfare Economics. **Journal of Economic Methodology**, v. 23, 1–25. 2016.
- JACQUEMET, N.; L'HARIDON, O. **Experimental Economics**: Method and Applications. Cambridge: Cambridge University Press, 2018.
- JACOBS, M.; MAZZUCATO, M. (ed.). **Rethinking Capitalism**: Economics and Policy for Sustainable and Inclusive Growth. Oxford: Wiley-Blackwell, 2016.
- JALLADEAU, J. Research Program versus Paradigm in the Development of Economics. **Journal of Economic Issues**, v. 12, p. 583-608, 1978.
- JANSSEN, M.; TAN, Y. Friedman's Permanent Income Hypothesis as an Example of Diagnostic Reasoning. **Economics and Philosophy**, v. 8, p. 23-50, 1992.
- JARVIE, I.; ZAMORA-BONILLA, J. **The SAGE Handbook of the Philosophy of Social Sciences**. New York: SAGE Publications, 2011.
- JENSEN, M.; MECKLING, W. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. **Journal of Financial Economics**, v. 3, 305–60. 1976. “
- JESPERSEN, J. **Macroeconomic Methodology**: A Post-Keynesian Perspective. Northampton: Edward Elgar, 2009.
- JEVONS, W. **The Theory of Political Economy**. 5. ed. London: MacMillan and Co, 1871.
- JOHNSON, J.; NOWAK, A.; ORMEROD, P.; ROSEWELL, B.; ZHAN, Y. (ed.). **Non-equilibrium Social Science and Policy**: Introduction and Essays on New and Changing Paradigms in Socio-economic Thinking. New York: Springer, 2017.

- KAGEL, J. H.; ROTH, A. E. (ed.). **The Handbook of Experimental Economics**. 2. ed. Princeton: Princeton University Press, 2008.
- KAGEL, J. H.; ROTH, A. E. (ed.). **The Handbook of Experimental Economics**. v. 2. Princeton: Princeton University Press, 2016.
- KAHNEMAN, D.; KNETSCH, J.; THALER, R. Fairness as a Constraint on Profit Seeking. **American Economic Review**, v. 76, p. 728-741, 1986.
- KALDOR, N. Welfare Propositions of Economics and Interpersonal Comparisons of Utility. **Economic Journal**, v49: 549–52. 1939
- KAMARCK, A. **Economics and the Real World**. Philadelphia: University of Pennsylvania Press, 1983.
- KAMARCK, A. **Economics for the Twenty-First Century**: The Economics of the Economist-Fox. Aldershot: Ashgate. 2001.
- KARLAN, D.; APPEL, J. **More Than Good Intentions**: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy. New York: Dutton, 2018.
- KARLAN, D.; APPEL, J. **Failing in the Field**: What We Can Learn When Field Research Goes Wrong, Princeton: Princeton University Press, 2016.
- KATOUZIAN, H. **Ideology and Method in Economics**. New York: New York University Press, 1980.
- KAUFMANN, F. On the Subject-Matter and Method of Economic Science. **Economica**, v. 13, p. 381-401, 1933.
- KAUFMANN, F. The Concept of Law in Economic Science. **Review of Economic Studies**, v. 1, p. 102-109, 1934.
- KAUFMANN, F. On the Postulates of Economic Theory. **Social Research**, v. 9, p. 379-395, 1942.
- KAUFMANN, F. **Methodology of the Social Sciences**. London: Oxford University Press, 1944.
- KEEN, S. **Debunking Economics**: The Naked Emperor of the Social Sciences. New York: St. Martin's Press, 2001.
- KEYNES, J. N. **The Scope and Method of Political Economy**. 4. ed. repr. New York: A. M. Kelley, 1955.
- KINCAID, H. **Philosophical Foundations of the Social Sciences**: Analyzing Controversies in Social Research. Cambridge: Cambridge University Press, 1996.
- KINCAID, H.; ROSS, D. (ed.). **Oxford Handbook of Philosophy of Economics**. Oxford: Oxford University Press, 2009.

- KINCAID, H. (ed.). **The Oxford Handbook of Philosophy of Social Science**. Oxford: Oxford University Press, 2012.
- KIRMAN, A. Who or What Does the Representative Agent Represent? **Journal of Economic Perspectives**, v. 6, p. 117-136, 1992.
- KIRZNER, I. **The Economic Point of View**. 2. ed. Kansas City: Sheed & Ward, 1976.
- KLAMER, A. **Conversations with Economists**: New Classical Economists and Opponents Speak Out on the Current Controversy in Macroeconomics. Totowa: Rowman and Allanheld, 1984.
- KLAMER, A.; COLANDER, D. **The Making of An Economist**. Boulder: Westview Press, 1990.
- KLAMER, A.; MCCLOSKEY, D.; SOLOW, R. (ed.). **The Consequences of Economic Rhetoric**. New York: Cambridge University Press, 1988.
- KLANT, J. **The Rules of the Game**. Cambridge: Cambridge University Press, 1984.
- KLANT, J. **The Nature of Economic Thought**: Essays in Economic Methodology. Cheltenham: Edward Elgar, 1994.
- KLAPPHOLZ, K. Value Judgments and Economics. **British Journal for the Philosophy of Science**, v. 15, p. 97-114, 1964.
- KLAPPHOLZ K.; AGASSI, J. Methodological Prescriptions in Economics. **Economica**, v. 26, p. 60-74, 1959.
- KNIGHT, F. Economics and Human Action. *In*: HAUSMAN, D. (ed.). **The Philosophy of Economics**: An Anthology. 3. ed. Cambridge: Cambridge University Press, 2008b, p. 111-118.
- KNIGHT, F. What is 'Truth' in Economics? **Journal of Political Economy**, v. 48, p. 1-32, 1940.
- KNIGHT, F. The Significance and Basic Postulates of Economic Theory: A Rejoinder. **Journal of Political Economy**, v. 49, p. 750-753, 1941.
- KNIGHT, F. Methodology in Economics. **Southern Economic Journal**, v. 27, p. 185-193; p. 273-282, 1961.
- KOOPMANS, T. **Three Essays on the State of Economic Science**. New York: McGraw-Hill, 1957.
- KOOPMANS, T. Economics Among the Sciences. **American Economic Review**, v. 69, p. 1-13, 1979.
- KOPPL, R. (ed.). **Explorations in Austrian Economics**. Bingley: JAI Press, 2008.
- KORNAI, J. **Anti-Equilibrium**: On Economic Systems Theory and the Tasks of Research, Amsterdam: North Holland. 1971.

- KOSLOWSKI, P. (ed.). **Economics and Philosophy**. Tübingen: J.C.B. Mohr, 1985.
- KREGEL, J. Economic Methodology in the Face of Uncertainty: The Modeling Methods of Keynes and the Post-Keynesians. **Economic Journal**, v. 86, p. 209-225, 1976.
- KREMER, M.; GLENNERSTER, R. **Small Changes, Big Results: Behavioral Economics at Work in Poor Countries**. Boston: Boston Review Press, 2011.
- KRUPP, S. (ed.). **The Structure of Economic Science**. Englewood Cliffs: Prentice-Hall, 1966.
- KUIPERS, T. (ed.). **What is Closer-to-the Truth? A Parade of Approaches to Truthlikeness**. Amsterdam: Rodopi, 1987.
- KUNIN, L.; WEAVER, F. On the Structure of Scientific Revolutions in Economics. **History of Political Economy**, v. 3, p. 391-397, 1971.
- KUORIKOSKI, J.; LEHTINEN, A. Incredible Worlds, Credible Results. **Erkenntnis**, v. 70, p. 119-131, 2009.
- KYDLAND, V.; PRESCOTT, E. The Econometrics of the General Equilibrium Approach to Business Cycles. **Scandinavian Journal of Economics**, v. 93, p. 161-178, 1991.
- KYDLAND, V.; PRESCOTT, E. The Computational Experiment: An Econometric Tool. **Journal of Economic Perspectives**, v. 10, p. 69-85, 1996.
- LACHMANN, L. Economics as a Social Science. **South African Journal of Economics**, v. 18, p. 233-241, 1950.
- LANGE, O. The Scope and Method of Economics. **Review of Economic Studies**, v. 13, p. 19-32, 1945.
- LATSIS, S. (ed.). **Method and Appraisal in Economics**. Cambridge: Cambridge University Press, 1976.
- LAVOIE, D. (ed.). **Economics and Hermeneutics**. London: Routledge, 1990.
- LAWSON, T. **Economics and Reality**. London: Routledge, 1997.
- LAWSON, T. **Essays on the Nature and State of Economic Theory**. London: Routledge, 2015.
- LAWSON, T.; PESARAN, H. **Keynes' Economics: Methodological Issues**. Beckenham. Kent: Croom Helm, 1985.
- LEAMER, E. Let's Take the Con Out of Econometrics. **American Economic Review**, v. 73, p. 31-43, 1983.
- LEAMER, E. **Vector Autoregressions for Causal Inference?** Delivered at 1984 Carnegie-Rochester Conference, 1984.

- LEBOWITZ, M. **The Socialist Imperative: From Gotha to Now**. New York: Monthly Review Press, 2015.
- LEE, F.; CRONIN, B. (ed.). **Handbook of Research Methods and Applications in Heterodox Economics**. Cheltenham: Elgar, 2016.
- LEHTINEN, J. K.; YLIKOSKI, P. (ed.). **Economics for Real: Uskali Mäki and the Place of Truth in Economics**. London: Routledge, 2012.
- LEIBENSTEIN, H. **Beyond Economic Man: A New Foundation for Economics**. Cambridge: Harvard University Press, 1976.
- LEIJONHUFVUD, A. **On Keynesian Economics and the Economics of Keynes**. Oxford: Oxford University Press, 1968.
- LEIJONHUFVUD, A. Life Among the Econ. **Western Economic Journal**, v. 11, p. 327-337, 1973.
- LEONARD, T. C. Reflection on Rules in Science: An Invisible-Hand Perspective. **Journal of Economic Methodology**, v. 9, p. 141-168, 2002.
- LEONTIEF, W. Theoretical Assumptions and Nonobserved Facts. **American Economic Review**, v. 61, p. 1-7, 1971.
- LESTER, R. A. Shortcomings of Marginal Analysis for Wage-Employment Problems. **American Economic Review**, v. 36, p. 62-82, 1946.
- LESTER, R. A. Marginal Costs, Minimum Wages, and Labor Markets. **American Economic Review**, v. 37, p. 135-148, 1947.
- LEVINE, A.; SOBER, E.; WRIGHT, E. Marxism and Methodological Individualism. **New Left Review**, v. 162, mar./abr., p. 67-84, 1987.
- LEVINE, A.; SOBER, E.; WRIGHT, E. **Reconstructing Marxism**. London: Verso, 1992.
- LEWIS, P. (ed.). **Transforming Economics**. London: Routledge, 2004.
- LICHTENSTEIN, S.; SLOVIC, P. (ed.). **The Construction of Preference**. New York: Cambridge University Press, 2006.
- LINSBICHLER, A. **Was Ludwig von Mises a Conventionalist? A New Analysis of the Epistemology of the Austrian School of Economics**. New York: Springer, 2017.
- LIPSEY, R.; LANCASTER, K. The General Theory of the Second Best. **Review of Economic Studies**, v. 24, p. 11-31, 1956-1957.
- LIST, J.; SAMEK, A. (ed.). **Field Experiments in Economics**. Northampton: Elgar, 2018.
- LITTLE, D. (ed.). **On the Reliability of Economic Models: Essays in the Philosophy of Economics**. Boston: Kluwer, 1993.
- LOASBY, B. **Choice, Complexity and Ignorance**. Cambridge: Cambridge University Press, 1976.

- LOASBY, B. **The Mind and Method of the Economist: A Critical Appraisal of Major Economists in the 20th Century.** Cheltenham: Edward Elgar, 1989.
- LOEWENSTEIN, G. **Exotic Preferences: Behavioral Economics and Human Motivation.** Oxford: Oxford University Press, 2007.
- LOEWENSTEIN, G.; RICK, S.; COHEN, J. D. Neuroeconomics. **Annual Review of Psychology**, v. 59, p. 647-672, 2008.
- LOWE, A. **On Economic Knowledge. Toward a Science of Political Economics.** New York: Harper & Row, 1965.
- LUCAS, R. Econometric Policy Evaluation: A Critique. **Journal of Monetary Economics**, v. 1, p. 19-46, 1976.
- MCCLELLAND, P. **Causal Explanation and Model Building in History, Economics and the New Economic History.** Ithaca: Cornell University Press, 1975.
- MCCLOSKEY, D. **The Rhetoric of Economics.** Madison: University of Wisconsin Press, 1985.
- MCCLOSKEY, D. **If You're So Smart: The Narrative of Economic Expertise.** Chicago: University of Chicago Press, 1992.
- MCCLOSKEY, D. **Truth and Persuasion in Economics.** Cambridge: Cambridge University Press, 1994.
- MCCLOSKEY, D. **How to be Human: Though an Economist University.** Ann Arbor: University of Michigan Press, 2000.
- MCCLOSKEY, D. N.; ZILIAK, S. T. **Measurement and Meaning in Economics: The Essential Deirdre McCloskey.** Northampton: Edward Elgar, 2003.
- MACHLUP, F. The Problem of Verification in Economics. **Southern Economic Journal**, v. 22, p. 1-21, 1955.
- MACHLUP, F. Operational Concepts and Mental Constructs in Model and Theory Formation. **Giornale Degli Economisti**, v. 19, p. 553-582, 1960.
- MACHLUP, F. **Essays on Economic Semantics.** Englewood Cliffs: Prentice-Hall, 1963.
- MACHLUP, F. Professor Samuelson on Theory and Realism. **American Economic Review**, v. 54, p. 733-736, 1964.
- MACHLUP, F. If Matter Could Talk. 1969a. In: MACHLUP, F. **Methodology of Economics and Other Social Sciences.** New York: Academic Press, 1978, p. 309-332.
- MACHLUP, F. Positive and Normative Economics. 1969b. In: MACHLUP, F. **Methodology of Economics and Other Social Sciences.** New York: Academic Press. 1978, p. 425-450.

- MACHLUP, F. **Methodology of Economics and Other Social Sciences**. New York: Academic Press. 1978.
- MACINTYRE, A. The Idea of a Social Science. **Proceedings of the Aristotelian Society Supplementar**, v. 41, p. 95-114, 1967.
- MCKENZIE, R. **The Limits of Economic Science**. Boston: Kluwer, 1983.
- MÄKI, U. How to Combine Rhetoric and Realism in the Methodology of Economics. **Economics and Philosophy**, v. 4, p. 89-109, 1988.
- MÄKI, U. Friedman and Realism. **Research in the History of Economic Thought and Methodology**, v. 10, 1990a.
- MÄKI, U. Mengerian Economics in Realist Perspective. **History of Political Economy**, v. 22, p. 289-310, 1990b.
- MÄKI, U. Scientific Realism and Austrian Explanation. **Review of Political Economy**, v. 2, p. 310-344, 1990c.
- MÄKI, U. On the Method of Isolation in Economics. *In*: DILWORTH, C. (ed.). **Intelligibility in Science**. Amsterdam: Rodopi, 1992, p. 317-351.
- MÄKI, U. Models Are Experiments. Experiments Are Models. **Journal of Economic Methodology**, v. 12, p. 303-315, 2005.
- MÄKI, U. On the method of isolation in economics. *In*: DAVIS, J.; DAVIS, J. B. (ed.). **Recent developments in economic methodology**, v. 3. Cheltenham: Edward Edgar, 2006, p. 3-37.
- MÄKI, U. **Realism and Economic Methodology**. London: Routledge, 2007.
- MÄKI, U. Missing the World. Models as Isolations and Credible Surrogate Systems. **Erkenntnis**, v. 70, p. 29-43, 2009a.
- MÄKI, U. Realistic Realism about Unrealistic Models. *In*: KINCAID, H.; ROSS, D. (ed.). **The Oxford Handbook of Philosophy of Economics**. New York: Oxford University Press, 2009b, p. 68-98.
- MÄKI, U. Models and the Locus of their Truth. **Synthese**, v. 180, p. 47-63, 2011.
- MÄKI, U. (ed.). **Fact and Fiction in Economics: Models, Realism and Social Construction**. Cambridge: Cambridge University Press, 1991.
- MÄKI, U. (ed.). **The Economic World View: Studies in the Ontology of Economics**. Cambridge: Cambridge University Press, 2001.
- MÄKI, U. (ed.). **The Methodology of Positive Economics: Reflections on the Milton Friedman Legacy**. Cambridge: Cambridge University Press, 2009c.
- MÄKI, U. (ed.). **Handbook of the Philosophy of Economics**. Amsterdam: Elsevier, 2012.
- MÄKI, U.; GUSTAFSSON, B.; KNUDSEN, C. (ed.). **Rationality, Institutions and Economic Methodology**. London: Routledge, 1993.

- MÄKI, U.; GABBAY, D. M.; THAGARD, P.; WOODS, J. (ed.). **The Rise and fall of Popper and Lakatos in Economics**. New York: Oxford University Press, 2007.
- MALINVAUD, E. **Lectures on Microeconomic Theory**. Amsterdam: North-Holland, 1972.
- MARCET, J. **Conversations on the Nature of Political Economy**. London: Transaction, 2009.
- MARCHIONNI, C.; VROMEN, J. (ed.). **Neuroeconomics: Hype or Hope?** London: Routledge, 2014.
- MARR, W.; RAJ, B. (ed.). **How Economists Explain: A Reader in Methodology**. Lanham: University Press of America, 1983.
- MARSCHAK, J. On Econometric Tools. **Synthese**, v. 20, p. 483-488, 1969.
- MAURO, C.; MIGUENS, S.; CADILHA, S. **Conversations on Human Action and Practical Rationality**. Newcastle: Cambridge Scholars, 2013.
- MARX, K. **Capital**. v. 1. New York: International Publishers, 1967.
- MARWELL, G.; AMES, R. Economists Free Ride. Does Anyone Else? Experiments on the Provision of Public Goods. IV. **Journal of Public Economics**, v. 15, p. 295-310, 1981.
- MAYER, T. **Truth Versus Precision in Economics**. Cheltenham: Edward Elgar, 1993.
- MEDEMA, S.; SAMUELS, W. (ed.). **Foundations of Research in Economics: How do Economists do Economics?** Cheltenham: Edward Elgar, 1996.
- MEEK, R. Value-Judgements in Economics. **British Journal for the Philosophy of Science**, v. 15, p. 89-96, 1964.
- MEIDINGER, C. **Science Économique: Questions de Méthode**. Paris: Vuibert, 1994.
- MELITZ, J. Friedman and Machlup on the Significance of Testing Economic Assumptions. **Journal of Political Economy**, v. 73, p. 37-60, 1965.
- MENGER, C. **Problems of Economics and Sociology**. Urbana: University of Illinois Press, 1963.
- MENGER, K. **Unexplored Dimensions: Karl Menger on Economics and Philosophy (1923–1938)**. Bingley: Emerald, 2009.
- MICOCCI, A. **A Historical Political Economy of Capitalism: After Metaphysics**. London: Routledge, 2016.
- MILL, J. S. **On the Definition of Political Economy and the Method of Investigation Proper to It**. Toronto: University of Toronto Press, 1967.
- MILL, J. S. **A System of Logic**. London: Longmans, Green & Co., 1949.
- MILL, J. S. **Principles of Political Economy**. 7. ed. New York: A. M. Kelley, 1976.

- MINFORD, P.; PEEL, D. **Rational Expectations and the New Macroeconomics**. Oxford: Martin Robertson & Co, 1983.
- MIROWSKI, P. **Against Mechanism**: Protecting Economics from Science. Totowa: Rowman and Littlefield, 1988.
- MIROWSKI, P. **More Heat Than Light**. Cambridge: Cambridge University Press, 1990.
- MIROWSKI, P. **Machine Dreams**: Economics Becomes a Cyborg Science. Cambridge: Cambridge University Press, 2002.
- MIROWSKI, P. **The Effortless Economy of Science?** Durham: Duke University Press, 2004.
- MIROWSKI, P. **Never Let a Serious Crisis Go to Waste**: How Neoliberalism Survived the Financial Meltdown. London: Verso, 2013.
- MIROWSKI, P. (ed.). **The Reconstruction of Economic Theory**. Boston: Kluwer, 1986.
- MIROWSKI, P.; SENT, E. (ed.). **Science Bought and Sold**: Essays in the Economics of Science. Chicago: University of Chicago Press, 2002.
- MISES, L. v. **Human Action. A Treatise on Economics**. New Haven: Yale University Press, 1949.
- MISES, L. v. **The Ultimate Foundation of Economic Science**: An Essay on Method. 2. ed. Kansas City: Sheed Andrews, 1978.
- MISES, L. v. **Epistemological Problems of Economics**. New York: New York University Press, 1981.
- MISHAN, E. **Cost Benefit Analysis**: An Introduction. New York: Praeger, 1971.
- MONGIN, P. La Controverse sur l'Entreprise (1940–1950) et la Formation de l'Irréalisme Méthodologique, **Economies et Sociétés**, v. 5, p. 91-151, 1986.
- MONGIN, P. The Full-Cost Controversy of the 1940s and 1950s: A Methodological Assessment. **History of Political Economy**, v. 24, p. 311-356, 1992.
- MORGAN, J. (ed.). **What is Neoclassical Economics? Debating the Origins, Meaning and Significance**. London: Routledge, 2016.
- MORGAN, M. Models, Stories, and the Economic World. **Journal of Economic Methodology**, v. 8, p. 361-384, 2001.
- MORGAN, M. Imagination and Imaging in Model Building. **Philosophy of Science**, v. 71, p. 753-766, 2004.
- MORGAN, M. **The World in the Model**: How Economists Work and Think. Cambridge: Cambridge University Press, 2012.
- MORGAN, M.; RUTHERFORD, M. (ed.). **From Interwar Pluralism to Postwar Neoclassicism**. Durham: Duke University Press, 1998.

- MUELLER, J. D. **Redeeming Economics**: Rediscovering the Missing Element. Wilmington: Intercollegiate Studies Institute, 2010.
- MULLAINATHAN, S.; SHAFIR, E. **Scarcity**: Why Having Too Little Means So Much. London: Times Books, 2013.
- MUSGRAVE, A. Unreal Assumptions in Economic Theory: The F-Twist Untwisted. **Kyklos**, v. 34, p. 377-387, 1981.
- MUTH, J. Rational Expectations and the Theory of Price Movements. **Econometrica**, v. 29, p. 315-335, 1961.
- MYRDAL, G. **The Political Element in the Development of Economic Thought**. Cambridge: Harvard University Press, 1955.
- NAGEL, E. Assumptions in Economic Theory. **American Economic Review Papers and Proceedings**, v. 53, p. 211-219, 1963.
- NELSON, A. New Individualistic Foundations for Economics. **Noûs**, v. 20, p. 469-490, 1986.
- NELSON, J. Feminism and Economics. **Journal of Economic Perspectives**, v. 9, p. 131-148, 1995.
- NELSON, J. **Feminism, Objectivity and Economics**. London: Routledge, 1996.
- NELSON, J. Economic Methodology and Feminist Critiques. **Journal of Economic Methodology**, v. 8, p. 93-97, 2001.
- NELL, L. **The Driving Force of the Collective**: Post-Austrian Theory in Response to Israel Kirzner. New York: Springer, 2017.
- NELL, G. (ed.). **Austrian Economic Perspectives on Individualism and Society**: Moving Beyond Methodological Individualism. New York: St. Martin's Press, 2014a.
- NELL, G. (ed.). **Austrian Theory and Economic Organization**: Reaching beyond Free Market Boundaries. New York: St. Martin's Press, 2014b.
- NELSON, R. **Economics As Religion**: From Samuelson to Chicago and Beyond. University Park: Pennsylvania State University Press, 2001.
- NELSON, R.; WINTER, S. **An Evolutionary Theory of Economic Change**. Cambridge: Harvard University Press, 1982.
- NEUBERG, L. **Conceptual Anomalies in Economics**. Cambridge: Cambridge University Press, 1988.
- NOWAK, L. **The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science**. Dordrecht: Reidel, 1980.
- OAKLEY, A. **Reconstructing Economic Theory**: The Problem of Human Agency. Cheltenham: Edward Elgar, 2002.

- O'BOYLE, E. **Personalist Economics: Moral Convictions, Economic Realities, and Social Action**. Boston: Kluwer, 1998.
- OCHANGCO, A. **Rationality in Economic Thought: Methodological Ideas on the History of Political Economy**. Cheltenham: Edward Elgar, 1999.
- OLIVER, A. (ed.). **Behavioural Public Policy**. Cambridge: Cambridge University Press, 2013.
- ORMEROD, P. **The Death of Economics**. New York: Wiley, 1997.
- ORMEROD, P. **Happiness, Economics and Public Policy**. London: Institute of Economic Affairs, 2007.
- O'SULLIVAN, P. **Economic Methodology and Freedom to Choose**. London: Allen & Unwin, 1987.
- PAPANDREOU, A. **Economics as a Science**. Chicago: Lippincott, 1958.
- PARETO, V. **Manual of Political Economy**. New York: A.M. Kelley, 1971.
- PARSONS, T. Some Reflections on The Nature and Significance of Economics. **Quarterly Journal of Economics**, v. 48, p. 511-545, 1934.
- PETERSON, M. (ed.). **The Prisoner's Dilemma**. New York: Cambridge University Press, 2015.
- PHEBY, J. **Methodology and Economics: A Critical Introduction**. London: Macmillan, 1988.
- PIKETTY, T. **Capital in the Twenty-First Century**. Cambridge: Harvard University Press, 2014.
- PITT, J. (ed.). **Philosophy in Economics**. Dordrecht: Reidel, 1981.
- PLOTT, C. R. Will Economics Become an Experimental Science? **Southern Economic Journal**, v. 57, 901–919. 1991.
- POPPER, K. La Rationalité et le Statut du Principe de Rationalité. In: CLASSEN, E. (ed.). **Les Fondements Philosophiques des Systèmes Économiques**. Paris: Paypot, 1967, p. 142-150.
- POPPER, K. The Logic of the Social Sciences. In: ADORNO, T. et. al. (ed.). **The Positivist Dispute in German Sociology**. New York: Harper, 1976. p. 87-104.
- POSNER, R. **Economic Analysis of Law**. Boston: Little, Brown & Co, 1972.
- PRATTEN, S. (ed.). **Social Ontology and Modern Economics**. London: Routledge, 2015.
- RABIN, M. Psychology and Economics. **Journal of Economic Literature**, v. 36, p. 11-46, 1998.
- RADCLIFF, B. **The Political Economy of Human Happiness**. Cambridge: Cambridge University Press, 2013.

- RAPPAPORT, S. **Models and Reality in Economics**. Cheltenham: Edward Elgar, 1998.
- REDER, M. **Economics: The Culture of a Controversial Science**. Chicago: University of Chicago Press, 1999.
- REDMAN, D. **Economic Methodology: A Bibliography with References to Works in the Philosophy of Science, 1860–1988**. New York: Greenwood Press, 1989.
- REDMAN, D. **Economics and the Philosophy of Science**. Oxford: Oxford University Press, 1990.
- REDMAN, D. **The Rise of Political Economy as a Science: Methodology and the Classical Economists**. Cambridge: MIT Press, 1997.
- REISS, J. **Error in Economics: Towards a More Evidence-Based Methodology**. London: Routledge, 2007.
- REISS, J. The Explanation Paradox. **Journal of Economic Methodology**, v. 19, p. 43-62, 2012.
- REISS, J. **The Philosophy of Economics: A Contemporary Introduction**. London: Routledge, 2013.
- REUTER, M.; MONTAG, C. (ed.). **Studies in Neuroscience, Psychology and Behavioral Economics**. New York: Springer, 2016.
- RICARDO, D. **On the Principles of Political Economy and Taxation**. Cambridge: Cambridge University Press, 1951.
- RICHTER, R. (ed.). **Essays on New Institutional Economics**. New York: Springer, 2015.
- ROBBINS, L. **An Essay on the Nature and Significance of Economic Science**. 3. ed. London: Macmillan, 1983.
- ROBINSON, J. **Economic Philosophy**. Chicago: Aldine, 1962.
- ROCHON, L.; ROSSI S. (ed.). **A Modern Guide to Rethinking Economics. New Directions in Post-Keynesian Economics**. Cheltenham: Elgar, 2017.
- RODRIK, D. **Economics Rules: The Rights and Wrongs of the Dismal Science**. New York: W.W. Norton, 2015.
- ROSCHER, W. **Geschichte der National-oekonomik in Deutschland**. Munich: R. Oldenbourg, 1874.
- ROSENBERG, A. **Microeconomic Laws: A Philosophical Analysis**. Pittsburgh: University of Pittsburgh Press, 1976.
- ROSENBERG, A. **Sociobiology and the Preemption of Social Science**. Baltimore: Johns-Hopkins University Press, 1980.
- ROSENBERG, A. Economics is too Important to Be Left to the Rhetoricians. **Economics and Philosophy**, v. 4, p. 129-149, 1988.

- ROSENBERG, A. **Economics - Mathematical Politics or Science of Diminishing Returns**. Chicago: University of Chicago Press, 1992.
- ROSS, D. **Economic Theory and Cognitive Science: Microexplanation**. Cambridge: MIT Press, 2005.
- ROSS, D. **Philosophy of Economics**. New York: St. Martin's Press, 2014.
- ROTH, A. Laboratory Experimentation in Economics: A Methodological Overview. **Economic Journal**, v. 98, p. 974-1031, 1988.
- ROTH, A. **Who Gets What—and Why: The New Economics of Matchmaking and Market Design**. New York: Houghton Mifflin Harcourt, 2015.
- ROTHBARD, M. In Defense of 'Extreme Apriorism'. **Southern Economic Journal**, v. 23, p. 314-320, 1957.
- ROTHSCHILD, K. **Ethics and Economic Theory**. Cheltenham: Edward Elgar, 1993.
- ROTWEIN, E. On The Methodology of Positive Economics. **Quarterly Journal of Economics**, v73: 554–75. 1959
- ROY, S. **Philosophy of Economics: On the Scope of Reason in Economic Inquiry**. London: Routledge, 1991.
- RUCCIO, D.; AMARIGLIO, J. **Postmodern Moments in Modern Economics**. Princeton: Princeton University Press, 2003.
- RUNDE, J. Assessing Causal Economic Explanations. **Oxford Economic Papers**, v. 50, p. 151-172, 1998.
- RUNDE, J.; MIZUHARA, S. (ed.). **The Philosophy of Keynes' Economics: Probability, Uncertainty and Convention**. London: Routledge, 2003.
- RUSSO, F. **Causality and Causal Modelling in the Social Sciences: Measuring Variations**. New York: Springer, 2009.
- RUSTICHINI, A. Neuroeconomics: Present and Future. **Games and Economic Behavior**, v. 52, p. 201-212, 2005.
- RUSTICHINI, A. Neuroeconomics: What have we found, and what should we search for? **Current Opinion in Neurobiology**, v. 19, p. 672-677, 2009.
- RUTHERFORD, M. **Institutions in Economics: The Old and New Institutionalism**. Cambridge: Cambridge University Press, 1994.
- SALANTI, A.; SCREPANTI, E. (ed.). **Pluralism in Economics: New Perspectives in History and Methodology**. Cheltenham: Edward Elgar, 1997.
- SAMUELS, W. **Erasing the Invisible Hand: Essays on an Elusive and Misused Concept in Economics**. Cambridge: Cambridge University Press, 2011.
- SAMUELS, W. (ed.). **The Methodology of Economic Thought: Critical Papers from the Journal of Economic Thought [Issues]**. New Brunswick: Transaction Books, 1980.

- SAMUELS, W. (ed.). **History and Methodology of Economics**. Greenwich: JAI Press, 1987.
- SAMUELS, W. (ed.). **Economics as Discourse**. Dordrecht: Kluwer, 1990.
- SAMUELSON, P. **Foundations of Economic Analysis**. Cambridge: Harvard University Press, 1947.
- SAMUELSON, P. Problems of Methodology - Discussion. **American Economic Review Papers and Proceedings**, v. 53, p. 232-236, 1963.
- SAMUELSON, P. Theory and Realism: A Reply. **American Economic Review**, v. 54, p. 736-740, 1964.
- SAMUELSON, P. Professor Samuelson on Theory and Realism: Reply. **American Economic Review**, v. 55, p. 1162-1172, 1965.
- SASSOWER, R. **Philosophy of Economics, A Critique of Demarcation**. Lanham: University Press of America, 1985.
- SCAZZIERI, R.; SEN, A.; ZAMAGNI, S. (ed.). **Markets, Money and Capital: Hicksian Economics for the Twenty-First Century**. Cambridge: Cambridge University Press, 2008.
- SCHLEFER, J. **The Assumptions Economists Make**. Cambridge: Harvard University Press, 2012.
- SCHMOLLER, G. **Zur Literatur-geschichte der Staats- und Sozialwissenschaften**. Leipzig: Duncker & Humblot, 1888.
- SCHMOLLER, G. **Über einige Grundfragen der Sozialpolitik und der Volkswirtschaftslehre**. Leipzig: Duncker & Humblot, 1898.
- SCHOEFFLER, S. **The Failures of Economics: A Diagnostic Study**. Cambridge: Harvard University Press, 1955.
- SCHRADER, D. **The Corporation as Anomaly**. Cambridge: Cambridge University Press, 1992.
- SCHUMPETER, J. **History of Economic Analysis**. New York: Oxford University Press, 1954.
- SCOTT, S. **Architectures of Economic Subjectivity: The Philosophical Foundations of the Subject in the History of Economic Thought**. New York: Routledge, 2013.
- SELIGMAN, B. On the Question of Operationalism: A Review Article. **American Economic Review**, v. 57, p. 146-161, 1967.
- SELIGMAN, B. The Impact of Positivism on Economic Thought. **History of Political Economy**, v. 1, p. 256-278, 1969.
- SEN, A.; WILLIAMS, B. (ed.). **Utilitarianism and Beyond**. Cambridge: Cambridge University Press, 1982.

- SENIOR, N. **Outline of the Science of Political Economy**. New York: A. M. Kelley, 1965.
- SENSAT, J. Methodological Individualism and Marxism. **Economics and Philosophy**, v. 4, p. 189-220, 1988.
- SENT, E. **The Evolving Rationality of Rational Expectations**. Cambridge: Cambridge University Press, 1998.
- SHACKLE, G. **Epistemics and Economics: A Critique of Economic Doctrines**. Cambridge: Cambridge University Press, 1972.
- SHARPES, D. K. **The Evolution of the Social Sciences**. Lanham: Lexington Books, 2009.
- SHRADER-FRECHETTE, K. **Science Policy, Ethics, and Economic Methodology: Some Problems of Technology Assessment and Environmental-Impact Analysis**. Dordrecht: D. Reidel, 1984.
- SIDGWICK, H. **The Scope and Method of Economic Science**. New York: A. M. Kelley, 1968.
- SIMON, H. Theories of Decision-Making in Economics and Behavioral Science. **American Economic Review**, v. 49, p. 253-283, 1959.
- SIMON, H. Problems of Methodology - Discussion. **American Economic Review Papers and Proceedings**, v. 53, p. 229-231, 1963.
- SIMON, H. **An Empirically Based Microeconomics**. Cambridge: Cambridge University Press, 1997.
- SIMS, C. Exogeneity and Causal Orderings in Macroeconomic Models. *In*: SIMS, C. (ed.). **New Methods in Business Cycle Research**. Minneapolis: Federal Reserve Bank, 1977, p. 23-43.
- SMITH, A. **An Inquiry into the Nature and Causes of the Wealth of Nations**. New York: Random House, 1937.
- SMYTH, R. (ed.). **Essays in Economic Method**. London: Duckworth, 1962.
- SOWELL, T. **Knowledge and Decisions**. New York: Basic Books, 1980.
- STANFIELD, R. Kuhnian Revolutions and the Keynesian Revolution. **Journal of Economic Issues**, v. 8, p. 97-109, 1974.
- STARMER, C. Experiments in Economics: should we trust the dismal scientists in white coats? **Journal of Economic Methodology**, v. 6, p. 1-30, 1999.
- STEWART, I. **Reasoning and Method in Economics. An Introduction to Economic Methodology**. London: McGraw-Hill, 1979.
- STIGLER, G. J. Professor Lester and the Marginalists. **American Economic Review**, v. 37, p. 154-157, 1947.
- STIGUM, B. **Econometrics and the Philosophy of Economics: Theory Data Confrontations in Economics**. Princeton: Princeton University Press, 2003.

- SUGDEN, R. Credible Worlds: The Status of Theoretical Models in Economics. **Journal of Economic Methodology**, v. 7, p. 1-31, 2000.
- SUGDEN, R. Credible Worlds, Capacities and Mechanisms. **Erkenntnis**, v. 70, p. 3-27, 2009.
- SUMMERS, L. The Scientific Illusion in Empirical Macroeconomics. **Scandinavian Journal of Economics**, v. 93, p. 129-148, 1991.
- SWEDBERG, R. **Economics and Sociology--Redefining Their Boundaries: Conversations with Economists and Sociologists**. Princeton: Princeton University Press, 1990.
- SWEDBERG, R. **Principles of Economic Sociology**. Princeton: Princeton University Press, 2007.
- THOMAS, R. Karl Popper and the Methodologists of Economics. **Cambridge Journal of Economics**, v. 41, p. 1143-1160, 2017.
- TITMUSS, R. **The Gift Relationship: From Human Blood to Social Policy**. New York: Random House, 1971.
- VEBLEN, T. Why Is Economics Not an Evolutionary Science? **Quarterly Journal of Economics**, v. 12, p. 373-397, 1898.
- VERCELLI, A. **Methodological Foundations of Macroeconomics: Keynes and Lucas**. Cambridge: Cambridge University Press, 1991.
- VERDON, M. **Keynes and the "Classics": A Study in Language, Epistemology and Mistaken Identities**. London: Routledge, 1996.
- VICKERS, D. **The Tyranny of the Market: A Critique of Theoretical Foundations**. Ann Arbor: University of Michigan Press, 1995.
- VROMEN, J. **Economic Evolution: An Inquiry into the Foundations of Institutional Economics**. London: Routledge, 1995.
- VROMEN, J. Advancing Evolutionary Explanations in Economics: The Limited Usefulness of Tinbergen's Four Questions Classification. In: KINCAID, H.; ROSS, D. (ed.). **Oxford Handbook of Philosophy of Economics**. Oxford: Oxford University Press, 2009, p. 337-368.
- VROMEN, J.; MARCHIONNI, C. (ed.). **Neuroeconomics**. London: Routledge, 2018.
- WARD, B. **What's Wrong with Economics?** New York: Basic Books. 1972.
- WEBER, M. Objectivity in Social Science and Social Policy. In: SHILS, E.; FINCH, H. (ed.). **The Methodology of the Social Sciences**. New York: Free Press, 1949, p. 49-112.
- WEIMANN, J.; KNABE, A. **Measuring Happiness: The Economics of Well-Being**. Cambridge: MIT Press, 2015.

- WEINTRAUB, E. R. **General Equilibrium Analysis: Studies in Appraisal**. Cambridge: Cambridge University Press, 1985.
- WEINTRAUB, E. R. **Stabilizing Dynamics: Constructing Economic Knowledge**. Cambridge: Cambridge University Press, 1991.
- WEINTRAUB, E. R. **How Economics Became a Mathematical Science**. Durham: Duke University Press, 2002.
- WIBLE, J. **The Economics of Science: Methodology and Epistemology as if Economics Really Mattered**. London: Routledge, 1998.
- WILBER, C.; HARRISON, R. The Methodological Basis of Institutional Economics: Pattern Model, Storytelling and Holism. **Journal of Economic Issues**, v. 12, p. 61-89, 1978.
- WILES, P.; ROUTH, G. (ed.). **What is Political Economy? Eight Perspectives**. Oxford: Basil Blackwell, 1984.
- WILSON, D.; KIRMAN, A. (ed.). **Complexity and Evolution: Toward a New Synthesis for Economics**. Cambridge: MIT Press, 2016.
- WINSTON, G.; TEICHGRAEBER, R. (ed.). **The Boundaries of Economics**. Cambridge: Cambridge University Press, 1988.
- WINTER, S. Economic 'Natural Selection' and the Theory of the Firm. **Yale Economic Essays**, v. 4, p. 255-272, 1962.
- WISEMAN, J. (ed.). **Beyond Positive Economics?** London: British Association for the Advancement of Science, 1983.
- WISMAN, J.; ROZANSKY, J. The Methodology of Institutionalism Revisited. **Journal of Economic Issues**, v. 25, p. 709-737, 1991.
- WITT, U. (ed.). **Recent Developments in Evolutionary Economics**. Cheltenham: Edward Elgar, 2008.
- WOLD, H. Causality and Econometrics. **Econometrica**, v. 22, p. 162-177, 1954.
- WOLPIN, K. **The Limits of Inference without Theory. Tjalling C. Koopmans Memorial Lectures**. Cambridge: MIT Press, 2013.
- WONG, S. **The Foundations of Paul Samuelson's Revealed Preference Theory**. London: Routledge, 1978.
- WORLAND, S. Radical Political Economy as a 'Scientific Revolution'. **Southern Economic Journal**, v. 39, p. 274-284, 1972.
- YEAGER, L. Methodenstreit, over Demand Curves. **Journal of Political Economy**, v. 68, p. 53-64, 1969.
- YUENGERT, A. **The Boundaries of Technique: Ordering Positive and Normative Concerns in Economic Research**. Lanham: Lexington Books, 2004.

- ZANINI, A. **Economic Philosophy**: Economic Foundations and Political Categories. New York: Lang, 2008.
- ZELLNER, A.; AIGNER, D. (ed.). Causality. **Journal of Econometrics**, v. 39, n. 1, 1988.
- ZILIAK, S.; MCCLOSKEY, D. **The Cult of Statistical Significance**: How the Standard Error Costs Us Jobs, Justice, and Lives. Ann Arbor: University of Michigan Press, 2008.

Ética e Economia

- ADLER, M. **Well-Being and Fair Distribution**: Beyond Cost-Benefit Analysis. New York: Oxford University Press, 2012.
- ADLER, M.; FLEURBAEY, M. (ed.). **The Oxford Handbook of Well-Being and Public Policy**. New York: Oxford University Press, 2016.
- ADLER, M.; POSNER, E. **New Foundations of Cost-Benefit Analysis**. Chicago: University of Chicago Press, 2006.
- ADLER, M.; POSNER, E. (ed.). **Cost-Benefit Analysis**: Legal, Economic and Philosophical Perspectives, Chicago: University of Chicago Press, 2000.
- ALLAIS, M. The Foundations of a Positive Theory of Choice involving Risk and a Criticism of the Postulates and Axioms of the American School. *In*: ALLAIS, M.; HAGEN, O. (ed.). **Expected Utility Hypotheses and the Allais Paradox**. Dordrecht: Reidel, 1979. p. 27-145.
- ANDERSON, E. The Ethical Limitations of the Market. **Economics and Philosophy**, v. 6, p. 179-206, 1990.
- ARNESON, R. Equality and Equal Opportunity for Welfare. **Philosophical Studies**, v. 56, p. 77-93, 1989.
- ARPALY, N.; SCHROEDER, T. **In Praise of Desire**. New York: Oxford University Press, 2014.
- ARROW, K. Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare. **Philosophy and Public Affairs**, v. 19, p. 158-194, 1990.
- ARROW, K. **Social Choice and Individual Values**. 2. ed. New York: Wiley, 1963.
- ARROW, K. Values and Collective Decision Making. *In*: HAHN, F.; HOLLIS, (ed.). **Philosophy and Economic Theory**. Oxford: Oxford University Press, 1979, p. 110-126.
- ARROW, K. Gifts and Exchanges. **Philosophy and Public Affairs**, v. 1, p. 343-362, 1972.

- ARROW, K. Some Ordinalist-Utilitarian Notes on Rawls Theory of Justice. **Journal of Philosophy**, v. 70, p. 245-263, 1973.
- ARROW, K. Extended Sympathy and the Possibility of Social Choice. **Philosophia**, v. 7, p. 223-237, 1978.
- ATKINSON, A. **Inequality: What Can Be Done?** Cambridge: Harvard University Press, 2015.
- AXELROD, R. **The Evolution of Cooperation**. New York: Basic Books, 1984.
- BAKER, C. The Ideology of the Economic Analysis of Law. **Philosophy and Public Affairs**, v. 5, p. 3-48, 1975.
- BALLET, J.; BAZIN, D.; DUBOIS, J.; MAHIEU, F. **Freedom, Responsibility and Economics of the Person**. London: Routledge, 2014.
- BAUMOL, W. **Superfairness: Applications and Theory**. Cambridge: MIT Press, 1986.
- BENTHAM, J. **An Introduction to the Principles of Morals and Legislation**. HARRISON, W. (ed.). Oxford: Basil Blackwell, 1967.
- BERGSON, A. A Reformulation of Certain Aspects of Welfare Economics. **Quarterly Journal of Economics**, v. 52, p. 30-34, 1938.
- BINMORE, K. **Playing Fair: Game Theory and the Social Contract**. Cambridge: MIT Press, 1994.
- BOADWAY, R. Cost-Benefit Analysis. In: ADLER, M.; FLEURBAEY, M. (ed.). **The Oxford Handbook of Well-Being and Public Policy**. New York: Oxford University Press, 2016, p. 47-81.
- BOARDMAN, A.; GREENBERG, D.; VINING, A.; WEIMER, D. **Cost-Benefit Analysis**. 4. ed. Englewood Cliffs: Prentice-Hall, 2010.
- BOULDING, K. Economics as a Moral Science. **American Economic Review**, v. 59, p. 1-12, 1969.
- BOWLES, S. **The New Economics of Inequality and Redistribution**. New York: Cambridge University Press, 2012.
- BOWLES, S.; GINTIS, H. A Political and Economic Case for the Democratic Enterprise. **Economics and Philosophy**, v. 9, p. 75-100, 1993.
- BOWLES, S.; GINTIS, H. **A Cooperative Species: Human Reciprocity and its Evolution**. Princeton: Princeton University Press, 2011.
- BRENNAN, G.; BUCHANAN, J. **The Reason of Rules: Constitutional Political Economy**. New York: Cambridge University Press, 1985.
- BROOME, J. Should Social Preferences Be Consistent? **Economics and Philosophy**, v. 5, p. 7-18, 1989.

- BROOME, J. **Weighing Goods**. Oxford: Basil Blackwell, 1991.
- BROOME, J. **Ethics Out of Economics**. Cambridge: Cambridge University Press, 1999.
- BROOME, J. **Weighing Lives**. Oxford: Oxford University Press, 2006.
- BROOME, J. **Climate Matters: Ethics in a Warming World**. New York: Norton, 2012.
- BUCHANAN, A. **Ethics, Efficiency, and the Market**. Totowa: Rowman & Allanheld, 1985.
- BUCHANAN, J. **The Limits of Liberty: Between Anarchy and Leviathan**. Chicago: University of Chicago Press, 1975.
- CARTER, I. **A Measure of Freedom**. Oxford: Oxford University Press, 1999.
- COASE, R. The Problem of Social Cost. **Journal of Law and Economics**, v. 3, p. 1-30, 1960.
- COHEN, G. A. On the Currency of Egalitarian Justice. **Ethics**, v. 99, p. 906-944, 1989.
- COLEMAN, J. Economics and the Law: A Critical Review of the Foundations of the Economic Approach to Law. **Ethics**, v. 94, p. 649-679, 1984.
- COLLARD, D. **Altruism and Economy: A Study in Non-selfish Economics**. New York: Oxford University Press, 1978.
- D'ASPREMONT, C.; GEVERS, L. Equity and the Informational Basis of Collective Choice. **Review of Economic Studies**, v. 44, p. 199-209, 1977.
- DEATON, A. **The Great Escape: Health, Wealth and the Origin of Inequality**. Princeton: Princeton University Press, 2013.
- DEBREU, G. **Theory of Value**. New York: Wiley, 1959.
- DRAKOPOULOS, S. **Values in Economic Theory**. Aldershot: Avebury, 1991.
- DWORKIN, G.; BERMANT, G.; BROWN, P. (ed.). **Markets and Morals**. Washington: Hemisphere Publishing, 1977.
- DWORKIN, R. What is Equality? Part 2: Equality of Resources. **Philosophy and Public Affairs**, v. 10, p. 283-345, 1981.
- EASTERLY, W. **The Tyranny of Experts: Economists, Dictators, and the Forgotten Rights of the Poor**. New York: Basic Books, 2013.
- EGE, R.; IGERSEIM, H. (ed.). **Freedom and Happiness in Economic Thought and Philosophy: From Clash to Reconciliation**. New York: Routledge, 2011.
- ELSTER, J.; HYLLAND, A. (ed.). **Foundations of Social Choice Theory**. New York: Cambridge University Press, 1986.
- ELSTER, J.; ROEMER, J. (ed.). **Interpersonal comparisons of well-being**. Cambridge: Cambridge University Press, 1991.
- EYAL, N.; HURST, S.; NORHEIM, O.; WIKLER, D. (ed.). **Inequalities in Health: Concepts, Measures and Ethics**. Oxford: Oxford University Press, 2013.

- FLEURBAEY, M. Equal Opportunity or Equal Social Outcome. **Economics and Philosophy**, v. 11, p. 25-56, 1995.
- FLEURBAEY, M. Equality of Resources Revisited. **Ethics**, v. 113, p. 82-105, 2002.
- FLEURBAEY, M. The Pazner-Schmeidler Social Ordering: A Defense. **Review of Economic Design**, v. 9, p. 145-166, 2005.
- FLEURBAEY, M. Social Choice and Just Institutions: New Perspectives. **Economics and Philosophy**, v. 23, p. 15-43, 2007.
- FLEURBAEY, M. **Fairness, Responsibility, and Welfare**. Oxford: Oxford University Press, 2008.
- FLEURBAEY, M. ; BLANCHET, D. **Beyond GDP: Measuring Welfare and Assessing Sustainability**. Oxford: Oxford University Press, 2013.
- FLEURBAEY, M.; MANIQUET, F. **A Theory of Fairness and Social Welfare**. Cambridge: Cambridge University Press, 2014.
- FRANK, R. **Passions Within Reason: The Strategic Role of the Emotions**. New York: W. W. Norton, 1988.
- FRANK, R. **Falling Behind: How Rising inequality Harms the Middle Class**. Berkeley: University of California Press, 2007.
- FRANK, R., T. Gilovich, and D. Regan, Does Studying Economics Inhibit Cooperation? **Journal of Economic Perspectives**, v. 7, p. 159-172, 1993.
- FANKHAUSER, S.; TOL, R.; PEARCE, D. The Aggregation of Climate Change Damages: A Welfare Theoretic Approach. **Environmental and Resource Economics**, v. 10, p. 249-266, 1997.
- FREY, B. **Happiness: A Revolution in Economics**. Cambridge: MIT Press, 2010.
- FREY, B.; STUTZER, A. **Happiness and Economics: How the Economy and Institutions Affect Human Well-Being**. Princeton: Princeton University Press, 2001.
- FRIEDMAN, M., **Capitalism and Freedom**. Chicago: University of Chicago Press, 1962.
- FRIEDMAN, M.; FRIEDMAN, R. **Free to Choose**. New York: Harcourt Brace Javanovich, 1980.
- GAERTNER, W.; SCHOKKAERT, E. **Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice**. Cambridge: Cambridge University Press, 2012.
- GAERTNER, W.; PATTANAIK, P.; SUZUMURA, K. Individual Rights Revisited. **Economica**, v. 59, p. 161-177, 1992.
- GAMBETTA, D. (ed.). **Trust: Making and Breaking Cooperative Relations**. Oxford: Basil Blackwell, 1988.

- GARDENFORS, P. Rights, Games and Social Choice. **Noûs**, v. 15, p. 341-356, 1981.
- GAUS, G. **The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World**. Cambridge: Cambridge University Press, 2011.
- GAUTHIER, D. **Morals by Agreement**. Oxford: Oxford University Press, 1986.
- GEORGE, D. **Preference Pollution: How Markets Creates Desires We Dislike**. Ann Arbor: University of Michigan Press, 2001.
- GIBBARD, A. A Pareto-Consistent Libertarian Claim. **Journal of Economic Theory**, v. 7, p. 388-410, 1974.
- GILBERT, M. Walking Together: A Paradigmatic Social Phenomenon. In: FRENCH, P.; UEHLING, T.; WETTSTEIN, H. (ed.). **Midwest Studies in Philosophy**. Notre Dame: University of Notre Dame Press, 1990, p. 1-14. (The Philosophy of the Human Sciences, v.15).
- GOTOH, R.; DUMOUCHEL, P. (ed.). **Against Injustice: The New Economics of Amartya Sen**, Cambridge: Cambridge University Press, 2009.
- GRANT, R. **Strings Attached: Untangling the Ethics of Incentives**. Princeton: Princeton University Press, 2012.
- GREAVES, H. Antiprioritarianism. **Utilitas**, v. 27, p. 1-42, 2015.
- GRETHER, D.; PLOTT, C. Economic Theory of Choice and the Preference Reversal Phenomenon. **American Economic Review**, v. 69, p. 623-638, 1979.
- HAMLIN, A., **Ethics, Economics, and the State**, New York: St. Martin's Press. 1986.
- HAMMOND, P. Ex-Post Optimality as a Dynamically Consistent Objective for Collective Choice Under Uncertainty. In: PATTANAİK, P.; SALLES, M. (ed.). **Social Choice and Welfare**. Amsterdam: North-Holland, 1983.
- HARBURGER, A. On the Use of Distributional Weights in Social Cost-Benefit Analysis. **Journal of Political Economy**, v. 86, p. s87-s120, 1978.
- HARDIN, R., **Collective Action**. Baltimore: Johns Hopkins University Press, 1982.
- HARRIS, R.; OLEWILER, N. The Welfare Economics of *Ex Post*, Optimality. **Economica**, v. 46, p. 137-147, 1979.
- HARSANYI, J. Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility. **Journal of Political Economy**, v. 63: p. 309-321, 1955.
- HARSANYI, J. Morality and the Theory of Rational Behavior. **Social Research**, v. 44; p. 39-62, 1977a.
- HAUSMAN, D. Hedonism and Welfare Economics. **Economics and Philosophy**, v. 26, p. 321-344, 2010.
- HAUSMAN, D.; MCPHERSON, M. **Economic Analysis, Moral Philosophy, and Public Policy**, Cambridge: Cambridge University Press, 2006.

- HAUSMAN, D.; MCPHERSON, M. "The Philosophical Foundations of Mainstream Normative Economics", in **The Philosophy of Economics: An Anthology**. 3. ed. Cambridge: Cambridge University Press, 2007, p. 226-250.
- HAUSMAN, D.; MCPHERSON, M. "Preference Satisfaction and Welfare Economics", **Economics and Philosophy**, v. 25, p. 1-25, 2009.
- HAUSMAN, D.; WALDREN, M. S. "Egalitarianism Reconsidered." **Journal of Moral Philosophy**, v. 8, p. 567-586, 2011.
- HAUSMAN, D.; WELCH, B. "To Nudge or Not to Nudge." **Journal of Political Philosophy**, v. 18, p. 123-36, 2010.
- HAUSMAN, D.; MCPHERSON, M.; SATZ, D. **Economic Analysis, Moral Philosophy, and Public Policy**. 3. ed. New York: Cambridge University Press, 2017.
- HAYEK, F. V. "The Moral Element in Free Enterprise." **Studies in Philosophy, Politics and Economics**, Chicago: University of Chicago Press, 1967, p. 229-236.
- HAYEK, F. V. **The Mirage of Social Justice**. Chicago: University of Chicago Press, 1976.
- HENNIPMAN, P. "Hicks, Robbins, and the Demise of Pigovian Welfare Economics: A Rectification and Amplification." **Southern Economic Journal**, v. 59, p. 88-97, 1992.
- HICKS, J. "The Foundations of Welfare Economics." **Economic Journal**, v. 49, p. 696-712, 1939.
- HIROSE, I. **Moral Aggregation**. Oxford: Oxford University Press, 2015.
- HIROSE, I.; OLSON, J. (ed.). **The Oxford Handbook of Value Theory**. New York: Oxford University Press, 2015.
- HIRSCH, F. **The Social Limits to Growth**. Cambridge: Harvard University Press, 1976.
- HOOK, S. (ed.). **Human Values and Economic Policy**. New York: New York University Press, 1967.
- KAHNEMAN, D. "Objective Happiness." In: KAHNEMAN, D.; DIENER, E.; SCHWARZ, N. (ed.). **Well-Being: Foundations of Hedonic Psychology**. New York: Russell Sage Foundation Press, 1999, p. 3-27.
- KAHNEMAN, D. "Evaluation by Moments: Past and Future." 2000a. In: KAHNEMAN, D.; TVERSKY, A. (ed.). **Choices, Values and Frames**. New York: Cambridge University Press and the Russell Sage Foundation, 2000, p. 693-708.
- KAHNEMAN, D. "Experienced Utility and Objective Happiness: A Moment-based Approach." 2000b. In: KAHNEMAN, D.; TVERSKY, A. (ed.). **Choices, Values and Frames**. New York: Cambridge University Press and the Russell Sage Foundation, 2000, p. 673-692.

- KAHNEMAN, D. **Thinking Fast and Slow**. New York: Farrar, Straus and Giroux, 2011.
- KAHNEMAN, D.; KRUEGER, A. Developments in the Measurement of Subjective Well Being. **Journal of Economic Perspectives**, v. 20, p. 3-24, 2006.
- KAHNEMAN, D.; SUGDEN, R. Experienced Utility as a Standard of Policy Evaluation. **Environmental & Resource Economics**, v. 32, p. 161-181, 2005.
- KAHNEMAN, D.; THALER, R. Utility Maximization and Experienced Utility. **Journal of Economic Perspectives**, v. 20, p. 221-234, 2006.
- KAHNEMAN, D.; TVERSKY, A. (ed.). **Choices, Values and Frames**. New York: Cambridge University Press and the Russell Sage Foundation, 2000.
- KALAI, E.; SMORODINSKY, M. Other Solutions to Nash's Bargaining Problem. **Econometrica**, v. 43, p. 513-518, 1975.
- KALDOR, N. Welfare Propositions of Economics and Interpersonal Comparisons of Utility. **Economic Journal**, v. 49, p. 549-552, 1939.
- KELMAN, S. **What Price Incentives?** Boston: Auburn House, 1981.
- KNIGHT, F. Economics and Human Action. In: **The Ethics of Competition, and other Essays**. New York and London: Harper & Brothers, 1935.
- KOLM, S.-C. **Justice et équité**. Paris: Editions du Centre National de la Recherche Scientifique, 1972.
- KRAUS, J.; COLEMAN, J. Morality and the Theory of Rational Choice. **Ethics**, v. 97, p. 715-749, 1987.
- LAYARD, R. **Happiness: Lessons from a New Science**. New York: Penguin, 2006.
- LE GRAND, J. **Equity and Choice**. London: Routledge, 1991.
- LE GRAND, J. **Motivation, Agency, and Public Policy: Of Knights and Knaves, Pawns and Queens**. Oxford: Oxford University Press, 2003.
- LE GRAND, J.; NEW, B. **Government Paternalism: Nanny State or Helpful Friend?** Princeton: Princeton University Press, 2015.
- LIPPERT-RASMUSSEN, K. **Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination**. Oxford: Oxford University Press, 2014.
- LITTLE, I. **A Critique of Welfare Economics**. 2. ed. Oxford: Oxford University Press, 1957.
- LOMASKY, L. **Persons, Rights and the Moral Community**. New York: Oxford University Press, 1987.
- MACCALLUM, G. Negative and Positive Freedom. **Philosophical Review**, v. 76, p. 312-334, 1967.
- MCCARTHY, D. Distributive equality. **Mind**, v. 124, p. 1045-1109, 2015.
- MCCARTHY, D. The Priority View. **Economics and Philosophy**, v. 33, p. 215-257, 2017.

- MACKAY, A. **Arrow's Theorem**: The Paradox of Social Choice. A Case Study in the Philosophy of Economics. New Haven: Yale University Press, 1980.
- MACKAY, A. Extended Sympathy and Interpersonal Utility Comparisons. **Journal of Philosophy**, v. 83, p. 305-322, 1986.
- MCKEAN, R. Economics of Trust, Altruism, and Corporate Responsibility. In: PHELPS, E. (ed.). **Altruism, Morality and Economic Theory**. New York: Russell Sage Foundation, 1975, p. 29-44.
- MANSBRIDGE, J. (ed.). **Beyond Self-interest**. Chicago: University of Chicago Press, 1990, p. 254-263.
- MARGOLIS, H. **Selfishness, Altruism and Rationality**. Cambridge: Cambridge University Press, 1982.
- MARWELL, G.; AMES, R. Economists Free Ride. Does Anyone Else? Experiments on the Provision of Public Goods. IV. **Journal of Public Economics**, v. 15, p. 295-310, 1981.
- MEADE, J. **Efficiency, Equality and the Ownership of Property**. London: George Allen & Unwin, 1964.
- MILL, J. S. **On Liberty**. Indianapolis: Hackett, 1978.
- MILLER, D. **Justice for Earthlings**: Essays in Political Philosophy. Cambridge: Cambridge University Press, 2012.
- MIROWSKI, P.; PLEHWE, D. (ed.). **The Road from Mont Pèlerin**: The Making of the Neoliberal Thought Collective. Cambridge: Harvard University Press, 2009.
- MISHAN, E. **Cost Benefit Analysis**: An Introduction. New York: Praeger, 1971.
- MONGIN, P. Consistent Bayesian Aggregation. **Econometrica**, v. 66, p. 313-351, 1995.
- MONGIN, P. Value Judgments and Value Neutrality in Economics. **Economica**, v. 73, p. 257-286, 2006.
- MORRIS, C. W. (ed.). **Amartya Sen. Contemporary Philosophy in Focus**. Cambridge: Cambridge University Press, 2010.
- NASH, J. The Bargaining Problem. **Econometrica**, v. 18, p. 155-162, 1950.
- NELSON, A. Economic Rationality and Morality. **Philosophy and Public Affairs**, v. 17, p. 149-166, 1988.
- NG, Y. **Welfare Economics**: Introduction and Development of Basic Concepts. Ed. rev. London: Macmillan, 1983.
- NUSSBAUM, M.; SEN, A. (ed.). **The Quality of Life**. Oxford: Clarendon Press, 1993.
- OKASHA, S.; BINMORE, K. **Evolution and Rationality**: Decisions, Cooperation, and Strategic Behaviour. Cambridge: Cambridge University Press, 2012.

- OKUN, A. **Equality and Efficiency: The Big Tradeoff**. Washington: Brookings Institution, 1975.
- PATTINAIK, P.; XU, Y. On Ranking Opportunity Sets in Terms of Freedom of Choice. **Rescherches Economiques de Louvain**, v. 56, p. 383-390, 1990.
- PAZNER, E.; SCHMEIDLER, D. A Difficulty in the Concept of Fairness. **Review of Economic Studies**, v. 41, p. 441-443, 1974.
- PETER, F.; SCHMID, H. B. (ed.). **Rationality and Commitment**. Oxford: Oxford University Press, 2007.
- PETTIT, P. Virtus Normativa: Rational Choice Perspectives. **Ethics**, v. 100, p. 725-755, 1990.
- PETTIT, P.; SUGDEN, R. The Backward Induction Paradox. **Journal of Philosophy**, v. 86, p. 169-182, 1989.
- POSNER, R. **Economic Analysis of Law**. Boston: Little, Brown & Co, 1972.
- REDER, M. The Place of Ethics in the Theory of Production. In: BOSKIN, M. (ed.). **Economics and Human Welfare: Essays in Honor of Tibor Scitovsky**. New York: Academic Press, 1979, p. 133-146.
- ROBERTSON, D. What Does the Economist Economize? In: **Economic Commentaries**. London: Staples Press, 1956, p. 147-155.
- ROEMER, J. Equality of Talent. **Economics and Philosophy**, v. 1, p. 151-188, 1985.
- ROEMER, J. The Mismatch of Bargaining Theory and Distributive Justice. **Ethics**, v. 97, p. 88-110, 1986a.
- ROEMER, J. Equality of Resources Implies Equality of Welfare. **Quarterly Journal of Economics**, v. 101, p. 751-784, 1986b.
- ROEMER, J. Egalitarianism, Responsibility, and Information. **Economics and Philosophy**, v. 3, p. 215-244, 1987.
- ROTHSCHILD, K. **Ethics and Economic Theory**. Cheltenham: Edward Elgar, 1993.
- SAINT-PAUL, G. **The Tyranny of Utility: Behavioral Social Science and the Rise of Paternalism**. Princeton: Princeton University Press, 2011.
- SAMUELSON, P. **Foundations of Economic Analysis**. Cambridge: Harvard University Press, 1947.
- SAMUELSON, P. Evaluation of Real National Income", **Oxford Economic Papers (New Series)**, v. 2, n. 1, p. 1-29, 1950.
- SANDEL, M. **What Money Can't Buy: The Moral Limits of Markets**. London: Allen Lane, 2012.
- SATZ, D. **Why Some Things Should Not be for Sale: The Moral Limits of Markets**. New York: Oxford University Press, 2011.

- SCANLON, T. Preference and Urgency. **Journal of Philosophy**, v. 72, p. 655-670, 1975.
- SCANLON, T. Equality of Resources and Equality of Welfare: A Forced Marriage? **Ethics**, 97: 111–18, 1986
- SCHOTTER, A. **The Economic Theory of Social Institutions**. Cambridge: Cambridge University Press, 1981.
- SCITOVSKY, T. A Note on Welfare Propositions in Economics. **Review of Economic Studies**, v. 9, p. 77-88, 1941.
- SEIDENFELD T.; KADANE, J.; SCHERVISH M. On the Shared Preferences of Two Bayesian Decision Makers. **Journal of Philosophy**, v. 86, p. 225-244, 1989.
- SEN, A. **Collective Welfare and Social Choice**. San Francisco: Holden-Day, 1970a.
- SEN, A. The Impossibility of a Paretian Liberal. **Journal of Political Economy**, v. 78, p. 152-157, 1970b.
- SEN, A. Liberty, Unanimity and Rights. **Economica**, v. 43, p. 217-245, 1976.
- SEN, A. *et. al.* **On Ethics and Economics**. Oxford: Blackwells, 1987a.
- SEN, A. The Standard of Living: Lecture I, Concepts and Critiques, 1987b. *In*: SEN, A. *et. al.* **On Ethics and Economics**. Oxford: Blackwells, 1987, p. 1-19.
- SEN, A. The Standard of Living: Lecture II, Lives and Capabilities, 1987c. *In*: SEN, A. *et. al.* **On Ethics and Economics**. Oxford: Blackwells, 1987, p. 20-38.
- SEN, A. Freedom of Choice: Concept and Content. **European Economic Review**, v. 32, p. 269-294, 1988.
- SEN, A. Welfare, Freedom and Social Choice: a Reply. **Rescherches Economiques de Louvain**, v. 56, 451-486, 1990.
- SEN, A. Welfare, Preference, and Freedom. **Journal of Econometrics**, v. 50, p. 15-29, 1991.
- SEN, A. **Inequality reexamined**. Cambridge: Harvard University Press, 1992.
- SEN, A. **Development as Freedom**. Oxford: Oxford University Press, 1999.
- SEN, A.; WILLIAMS, B. (ed.). **Utilitarianism and Beyond**. Cambridge: Cambridge University Press, 1982.
- SHIFFRIN, S. Paternalism, Unconscionability Doctrine, and Accommodation. **Philosophy & Public Affairs**, v. 29, p. 205-550, 2000.
- SIDGWICK, H. **The Methods of Ethics**, 6. ed. London: Macmillan, 1901.
- STEEL, D. **Philosophy of the Precautionary Principle**. Cambridge: Cambridge University Press, 2015.
- STREETEN, P. Appendix: Recent Controversies. *In*: MYRDAL, G. **The Political Element in the Development of Economic Theory**. London: Routledge and Kegan Paul, 1953, p. 208-217.

- SUGDEN, R. Liberty, Preference and Choice. **Economics and Philosophy**, v. 2, p. 213-231, 1985.
- SUGDEN, R. **The Economics of Rights, Co-operation and Welfare**. New York: Blackwell, 1986.
- SUGDEN, R. Spontaneous Order. **Journal of Economic Perspectives**, v. 3: 85-97. 1989.
- SUGDEN, R. Contractarianism and Norms. **Ethics**, v. 100, p. 768-786, 1990.
- SUGDEN, R. **The Community of Advantage: A Behavioural Economist's Defence of the Market**. Oxford: Oxford University Press, 2018.
- SUGDEN, R.; WILLIAMS, A. **The Principles of Practical Cost-benefit Analysis**. New York: Oxford University Press, 1978.
- TAYLOR, M. **The Possibility of Cooperation**. New York: Cambridge University Press, 1987.
- THALER, R.; SUNSTEIN, C. **Nudge: Improving Decisions about Health, Wealth, and Happiness**. New Haven: Yale University Press, 2008.
- TITMUSS, R. **The Gift Relationship: From Human Blood to Social Policy**. New York: Random House, 1971.
- TVERSKY, A.; THALER, R. Preference Reversals. **Journal of Economic Perspectives**, v. 4, p. 201-211, 1990.
- VAN PARIJS, P. The Second Marriage of Justice and Efficiency. **Journal of Social Policy**, v. 19, p. 1-25, 1990.
- VARIAN, H. Equity, Envy and Efficiency. **Journal of Economic Theory**, v. 9, p. 63-91, 1974.
- VARIAN, H. Distributive Justice, Welfare Economics and the Theory of Fairness. **Philosophy and Public Affairs**, v. 4, p. 223-247, 1975.
- VARIAN, H. Dworkin on Equality of Resources. **Economics and Philosophy**, v. 1, p. 110-127, 1985.
- VICKERS, D. **Economics and Ethics: An Introduction to Theory, Institutions, and Policy**. London: Praeger, 1997.
- VICKREY, W. Utility, Strategy, and Social Decision Rules. **Quarterly Journal of Economics**, v. 74, p. 507-535, 1960.
- WENAR, L. **Blood Oil: Tyrants, Violence and the Rules that Run the World**. Oxford: Oxford University Press, 2016.
- WEYMARK, J. A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism. In: ELSTER, J.; ROEMER, J. (ed.). **Interpersonal comparisons of well-being**. Cambridge: Cambridge University Press, 1991, p. 255-320.

YAARI, M.; BAR-HILLEL, M. On Dividing Justly. **Social Choice and Welfare**, v. 1, p. 1-24, 1984.

Racionalidade

ALLAIS, M.; HAGEN, O. (ed.). **Expected Utility Hypotheses and the Allais Paradox**. Dordrecht: Reidel, 1979.

BARBERÀ, S.; HAMMOND, P.; SEIDL, C. **Handbook of Utility Theory**: volume 1 Principles. Dordrecht: Kluwe, 1999.

BICCHIERI, C. **Rational and Coordination**. Cambridge: Cambridge University Press, 1993.

BINMORE, K. Modeling Rational Players. **Economics and Philosophy**, v. 3, p. 179-214; v. 4, p. 9-56, 1987/1988.

BINMORE, K. **Fun and games**. New York: D.C Heath, 1992.

BONNANO, G. The Logic of Rational Play in Games of Perfect Information. **Economics and Philosophy**, v. 7, p. 37-65, 1991.

BROOME, J. Utility. **Economics and Philosophy**, v. 7, p. 1-12, 1991.

BUCHAK, L. **Risk and Rationality**. New York: Oxford University Press, 2013.

DENNIS, K. (ed.). **Rationality in Economics**: Alternative Perspectives. Boston: Kluwer, 1998.

EELS, E. **Rational Decision and Causality**. Cambridge: Cambridge University Press, 1982.

ELLSBERG, D. Classic and Current Notions of Measurable Utility. **Economic Journal**, v. 64, p. 528-556, 1954.

ELSTER, J. **Ulysses and the Sirens**: Studies in Rationality and Irrationality. Cambridge: Cambridge University Press, 1979.

ELSTER, J. **Sour Grapes**: Studies in the Subversion of Rationality. Cambridge: Cambridge University Press, 1983.

FRIEDMAN, M.; SAVAGE, L. The Utility Analysis of Choices Involving Risk. **Journal of Political Economy**, v. 56, p. 279-304, 1948.

FRIEDMAN, M.; SAVAGE, L. The Expected-Utility Hypothesis and the Measurability of Utility. **Journal of Political Economy**, v. 60, 463-474, 1952.

GERRARD, B. **The Economics of Rationality**. London: Routledge, 1993.

GILBOA, I. **Rational Choice**. Cambridge: MIT Press, 2010.

- GILBOA, I.; SAMUELSON, L.; SCHMEIDLER, D. **Analogies and Theories: Formal Models of Reasoning**. New York: Oxford University Press, 2015.
- GILBOA, I.; SCHMEIDLER, D. **A Theory of Case-Based Decisions**. New York: Cambridge University Press Press, 2001.
- GREETHER, D.; PLOTT, C. Economic Theory of Choice and the Preference Reversal Phenomenon. **American Economic Review**, v. 69, p. 623-638, 1979.
- HARGREAVES-HEAP, S. **Rationality in Economics**. Oxford: Blackwell, 1989.
- HARSANYI, J. **Rational Behavior and Bargaining Equilibrium in Games and Social Situations**. Cambridge: Cambridge University Press, 1977b.
- HAUSMAN, D. **Preference, Value, Choice and Welfare**. Cambridge: Cambridge University Press, 2012.
- HERNSTEIN, I.; MILNOR, J. An Axiomatic Approach to Measurable Utility. **Econometrica**, v. 21, p. 291-297, 1953.
- HOUTTHAKER, H. Revealed Preference and the Utility Function. **Economica**, v. 17, p. 159-174, 1950.
- HOWSON, C.; URBACH, P. **Scientific Reasoning: The Bayesian Approach**. LaSalle: Open Court, 1989.
- JEFFREY, R. **The Logic of Decision**. 2. ed. Chicago: University of Chicago Press, 1983.
- KAHNEMAN, D.; TVERSKY, A. Prospect Theory: An Analysis of Decision Making under Risk. **Econometrica**, v. 47, p. 263-291, 1979.
- KREPS, D.; MILGROM, P.; ROBERTS, J.; WILSON, R. Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. **Journal of Economic Theory**, v. 27, p. 245-252, 1982.
- LEVI, I. **The Enterprise of Knowledge**. Cambridge: MIT Press, 1980.
- LEVI, I. The Paradoxes of Allais and Ellsberg. **Economics and Philosophy**, v. 2, p. 23-53, 1986.
- LEVI, I. Pareto Unanimity and Consensus. **Journal of Philosophy**, v. 89, 481-492, 1990.
- LICHTENSTEIN, S.; SLOVIC, P. Reversals of Preference Between Bids and Choices in Gambling Decisions. **Journal of Experimental Psychology**, v. 89, p. 46-55, 1971.
- LIST, C.; PETTIT, P. Aggregating Sets of Judgments: An Impossibility Result. **Economics and Philosophy**, v. 18, p. 89-110, 2002.
- LOOMES, G.; SUGDEN, R. Regret Theory: an Alternative Theory of Rational Choice under Uncertainty. **Economic Journal**, v. 92, p. 805-824, 1982.
- LUCE, R.; RAIFFA, H. **Games and Decisions**. New York: Wiley, 1957.

- MCCLENNEN, E. Sure Thing Doubts. *In*: STIGUM, B.; WENSTOP, F. (ed.). **Foundations of Utility and Risk Theory with Applications**. Dordrecht: Reidel, 1983.
- MCCLENNEN, E. **Rationality and Dynamic Choice**: Foundational Explorations. Cambridge: Cambridge University Press, 1990.
- MACHINA, M. Choice under Uncertainty: Problems Solved and Unsolved. **Journal of Economic Perspectives**, v. 1, p. 121-154, 1987.
- PAUL, L. A. **Epistemic Transformation and Rational Choice**: Transformative Experience. Oxford: Oxford University Press, 2014.
- PETTIT, P.; SUGDEN, R. The Backward Induction Paradox. **Journal of Philosophy**, v. 86, p. 169-182, 1989.
- RAMSEY, F. Truth and Probability. *In*: BRAITHWAITE, R. (ed.). **The Foundations of Mathematics and other Logical Essays**. London: Routledge & Kegan Paul, 1926, p. 156-198.
- RESNIK, M. **Choices**: An Introduction to Decision Theory. Minneapolis: University of Minnesota Press, 1987.
- ROTH, A.; MALOUF, M. Game Theoretical Models and the Role of Information in Bargaining. **Psychological Review**, v. 86, p. 574-594, 1979.
- SAVAGE, L. **The Foundations of Statistics**. New York: Dover, 1972.
- SCHICK, F. Money Pumps and Dutch Bookies. **Journal of Philosophy**, v. 83, p. 112-119, 1986.
- SEN, A. **Collective Welfare and Social Choice**. San Francisco: Holden-Day, 1970.
- SEN, A. Choice Functions and Revealed Preference. **Review of Economic Studies**, v. 38, p. 307-317, 1971.
- SEN, A. Behaviour and the Concept of Preference. **Economica**, v. 40, p. 241-259, 1973.
- SEN, A. Rational Fools. *In*: HAHN, F.; HOLLIS, (ed.). **Philosophy and Economic Theory**. Oxford: Oxford University Press, 1979, p. 87-109.
- SIMON, H. From Substantive to Procedural Rationality. *In*: LATSIS, S. (ed.). **Method and Appraisal in Economics**. Cambridge: Cambridge University Press, 1976, p. 129-148.
- STONE, P. **The Luck of the Draw**: The Role of Lotteries in Decision Making. Oxford: Oxford University Press, 2011.
- SUGDEN, R. New Developments in the Theory of Choice Under Uncertainty. **Bulletin of Economic Research**, v. 38, p. 1-24, 1986.
- VICKREY, W. Measuring Marginal Utility by Reactions to Risk. **Econometrica**, v. 13, p. 319-333, 1945.

- VON NEUMANN, J.; MORGENSTERN, O. **Theory of Games and Economic Behavior**. 2. ed. Princeton: Princeton University Press, 1947.
- YOUNG, P. **Individual Strategy and Social Structure**. Princeton: Princeton University Press, 1998.

Outros trabalhos citados

- BACKHOUSE, R. **The Ordinary Business of Life**. Princeton: Princeton University Press, 2002.
- BECKER, G. **A Treatise on the Family**. Cambridge: Harvard University Press, 1981.
- BENTHAM, J. **An Introduction to the Principles of Morals and Legislation**. Oxford: Basil Blackwell, 1967.
- BHASKAR, R. **A Realist Theory of Science**. Leeds: Leeds Press, 1975.
- BOYD, R. The Current Status of Scientific Realism. In: LEPLIN, J. (ed.). **Scientific Realism**. Berkeley: University of California Press, 1984, p. 41-82.
- CANTILLON, R. **Essai sur la nature du commerce en général**. Paris: Institut national d'études démographiques, 1952.
- CARTWRIGHT, N. **How the Laws of Physics Lie**. Oxford: Clarendon Press, 1983.
- CARTWRIGHT, N. **Nature's Capacities and their Measurement**. Oxford: Clarendon Press, 1989.
- CARTWRIGHT, N. **Hunting Causes and Using Them**: Approaches in Philosophy and Economics. Cambridge: Cambridge University Press, 2007.
- CARTWRIGHT, N.; HARDIE, J. **Evidence-Based Policy**: A Practical Guide to Doing it Better. Oxford: Oxford University Press, 2013.
- CHAKRAVARTTY, A. **A Metaphysics for Scientific Realism**: Knowing the Unobservable. Cambridge: Cambridge University Press, 2010.
- COASE, R. The Nature of the Firm. **Economica**, v. 4, p. 386-405, 1937.
- DAVIDSON, D. Actions, Reasons and Causes. **Journal of Philosophy**, v. 60, p. 685-700, 1963.
- DICKEN, P. **A Critical Introduction to Scientific Realism**. London: Bloomsbury Academic Press, 2016.
- DUHEM, P. **The Aim and Structure of Scientific Theories**. Princeton: Princeton University Press, 1954.
- FISH, S. **Is There a Text in This Class? The Authority of Interpretive Communities**. Cambridge: Harvard University Press, 1980.

- FRIGG, R. Fiction and Scientific Representation. *In*: FRIGG, R.; HUNTER, M. C. (ed.). **Beyond Mimesis and Convention**. New York: Springer, 2010.
- GODFREY-SMITH, P. The Strategy of Model-Based Science. **Biology and Philosophy**, v. 21, p. 725-740, 2006.
- KUHN, T. **The Structure of Scientific Revolutions**. 2. ed. Chicago: University of Chicago Press, 1970.
- LAKATOS, I. Falsification and the Methodology of Scientific Research Programmes. *In*: LAKATOS, I.; MUSGRAVE, A. (ed.). **Criticism and the Growth of Knowledge**. Cambridge: Cambridge University Press, 1970, p. 91-196.
- LAKATOS, I. Popper on Demarcation and Induction. *In*: SCHLIPP, P. (ed.). **The Philosophy of Karl Popper**. LaSalle: Open Court, 1974, p. 241-273.
- LAKATOS, I.; MUSGRAVE, A. (ed.). **Criticism and the Growth of Knowledge**. Cambridge: Cambridge University Press, 1970.
- MORGAN, M.; MORRISON, M. (ed.). **Models as Mediators: Perspectives on Natural and Social Science**. Cambridge: Cambridge University Press, 1999.
- MORGENBESSER, S. **Theories and Schemata in the Social Sciences**. Pennsylvania: University of Pennsylvania, 1956.
- MORISHIMA, M. **Marx's Economics: A Dual Theory of Value and Growth**. Cambridge: Cambridge University Press, 1973.
- NIINILUOTO, I. **Critical Scientific Realism**. Oxford: Oxford University Press, 2002.
- NORTH, D. **Institutions, Institutional Change and Economic Performance**. Cambridge: Cambridge University Press, 1990.
- NOZICK, R. **Anarchy, State and Utopia**. New York: Basic Books, 1974.
- NUSSBAUM, M. **Women and Economic Development: The Capabilities Approach**. Cambridge: Cambridge University Press, 2000.
- OCHS, J.; ROTH, A. An Experimental Study of Sequential Bargaining. **American Economic Review**, v. 79, p. 355-384, 1989.
- PASINETTI, L. **Structural Change and Economic Growth: A Theoretical Essay on the Dynamics of the Wealth of Nations**. Cambridge: Cambridge University Press, 1981.
- PEARL, J. **Causality**. Cambridge: Cambridge University Press, 2000.
- POPPER, K. **The Logic of Scientific Discovery**. ed. rev. London: Hutchinson & Co, 1968.
- POPPER, K. **Conjectures and Refutations; The Growth of Scientific Knowledge**. 3. ed. London: Routledge & Kegan-Paul, 1969.
- PSILLOS, S. **Scientific Realism: How Science Tracks Truth**. London: Routledge, 1999.

- PUTNAM, H. The Analytic and the Synthetic. *In: Mind, Language and Reality: Philosophical Papers*. Cambridge University Press, 1975, p. 33-69.
- QUINE, W. V. O. Two Dogmas of Empiricism. *In: From a Logical Point of View*. Cambridge: Harvard University Press, 1953, p. 20-46.
- ROEMER, J. **Analytical Foundations of Marxian Economic Theory**. Cambridge: Cambridge University Press, 1981.
- ROEMER, J. **A General Theory of Exploitation and Class**. Cambridge: Harvard University Press, 1982.
- RONCAGLIA, A. **Sraffa and the Theory of Prices**. Chicester: John Wiley, 1978.
- SMITH, V.L. **Papers in Experimental Economics**. Cambridge, Cambridge University Press, 1991.
- SNEED, J. **The Logical Structure of Mathematical Physics**. Dordrecht: Reidel, 1971.
- SPIRITES, P.; GLYMOUR, C.; SCHEINES, R. **Causation, Prediction and Search**. 2. ed. Cambridge: MIT Press, 2001.
- SRAFFA, P. **Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory**. Cambridge: Cambridge University Press, 1960.
- STEGMUELLER, W. **The Structure and Dynamics of Theories**. New York: Springer-Verlag, 1976.
- STEGMUELLER, W. **The Structuralist View of Theories**. New York: Springer-Verlag, 1979.
- STEGMUELLER, W.; BALZER, W.; SPOHN, W. (ed.). **Philosophy of Economics**. New York: Springer-Verlag, 1982. (Proceedings, Munich, July 1981).
- SWEDBERG, R. **The Art of Social Theory**. Princeton: Princeton University Press, 2014.
- TUOMELA, R. **Social Ontology: Collective Intentionality and Group Agents**. Oxford: Oxford University Press, 2016.
- VAN FRAASSEN, B. **The Scientific Image**. Oxford: Oxford University Press, 1980.
- VON WRIGHT, G. H. **Explanation and Understanding**. Ithaca: Cornell University Press, 1971.
- WATKINS, J. **Science and Scepticism**. Princeton: Princeton University Press, 1984.
- WEISBERG, M. "Who Is a Modeler?" **British Journal for the Philosophy of Science**, v. 58, p. 207-233, 2007.
- WILLIAMS, B. Internal and External Reasons. *In: Moral Luck*. Cambridge: Cambridge University Press, 1981, p. 101-113.
- WILLIAMSON, O. **The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting**. New York: Free Press, 1985.

- WINCH, P. **The Idea of a Social Science**. London: Routledge, 1958.
- WOODWARD, J. **Making Things Happen**. New York: Oxford University Press, 2003.
- WORRALL, J. Why There's No Cause to Randomize. **The British Journal for the Philosophy of Science**, v. 58, n. 3, p. 451-488, 2007.

Outros recursos na Internet:

Websites

- Economics and Philosophy**. Disponível em: <<https://www.cambridge.org/core/journals/economics-and-philosophy>>. Acesso em: 16 jan. 2022.
- History of Economics Society**. Disponível em: https://eh.net/book_reviews/health-and-labor-force-participation-over-the-life-cycle-evidence-from-the-past/. Acesso em: 16 jan. 2022.
- Hoover Chair in Economic and Social Ethics**. Disponível em: <https://uclouvain.be/chaire-hoover>. Acesso em: 16 jan. 2022.

Blogs

Há, hoje, muitos *blogs* sérios de proeminentes economistas. Embora não sejam predominantemente voltados para metodologia, nem exclusivamente voltados à economia, esses *blogs* mostram economistas discutindo uns com os outros, respondendo aos eventos do momento e formulando e reformulando seus posicionamentos. Os seguintes *blogs* são especialmente interessantes, mas há muitos outros:

- Brad DeLong Grasping Reality with Both Invisible Hands: Fair, Balanced, and Reality-Based: A Semi-Daily Journal**. Disponível em: <https://delong.typepad.com/>. Acesso em: 16 jan. 2022.
- Paul Krugman, The Conscience of a Liberal**. Disponível em: <https://krugman.blogs.nytimes.com/>. Acesso em: 16 jan. 2022.

- Marginal Revolution:** Small Steps Toward a Much Better World (Tyler Cowen and Alex Tabbarok). Disponível em: <http://marginalrevolution.com/>. Acesso em: 16 jan. 2022.
- Mainly Macro:** Comment on Macroeconomic Issues (Simon Wren-Lewis). Disponível em: <https://mainlymacro.blogspot.com/>. Acesso em: 16 jan. 2022.
- The Grumpy Economist (John Cochrane's Blog).** Disponível em: <https://johnhcochrane.blogspot.com/>. Acesso em: 16 jan. 2022.
- Dani Rodrik's Weblog:** Unconventional thoughts on economic development and globalization. Disponível em: <https://rodrik.typepad.com/>. Acesso em: 16 jan. 2022.
- Beat the Press** (Dean Baker). Disponível em: <https://www.cepr.net/beat-the-press-is-it-time-for-a-name-change/>. Acesso em: 16 jan. 2022.

(II) Teoria da Decisão*

Katie Steele (ANU) e H. Orri Stefánsson (Stockholm University)

Tradução: Ramiro de Ávila Peres (BCB)

Revisão: Mariana Kuhn de Oliveira (CEBRAP)

A teoria da decisão se preocupa com o raciocínio subjacente às escolhas de um agente, seja uma escolha banal entre pegar o ônibus ou pegar um táxi, ou uma escolha mais abrangente sobre se deve ou não seguir uma carreira política exigente (note que “agente” aqui significa uma entidade, geralmente uma pessoa física, que é capaz de deliberação e ação). O pensamento padrão é que o que um agente faz em qualquer ocasião é completamente determinado por suas crenças e desejos/valores, mas isso não é incontroverso, como será observado abaixo. Em qualquer caso, a teoria da decisão é tanto uma teoria de crenças, desejos e outras atitudes relevantes quanto uma teoria da escolha; o que importa é como essas várias atitudes (chamadas de “atitudes de preferência”) coexistem entre si.

*STEELE, K.; STEFÁNSSON, H. O. Decision Theory. In: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Winter Edition. Stanford, CA: The Metaphysics Research Lab, 2020. Disponível em: <https://plato.stanford.edu/archives/win2020/entries/decision-theory/>.

The following is the translation of the entry on “Decision Theory” by Katie Steele and H. Orri Stefánsson in the Stanford Encyclopedia of Philosophy. The translation follows the version of the entry in the SEP’s archives at <https://plato.stanford.edu/archives/win2020/entries/decision-theory/>. This translated version may differ from the current version of the entry, which may have been updated since the time of this translation. The current version is located at a <https://plato.stanford.edu/entries/ethics-business/>. We’d like to thank the Editors of the Stanford Encyclopedia of Philosophy, mainly Prof. Dr. Edward Zalta, for granting permission to translate and to publish this entry.

O foco deste verbete é a teoria da decisão normativa. Ou seja, a principal questão de interesse é que critérios as atitudes de preferência de um agente **devem** satisfazer em **quaisquer** circunstâncias **genéricas**. Isso equivale a uma abordagem mínima de **racionalidade**, que, dada a situação em questão, deixa de lado questões mais substanciais sobre preferências apropriadas, e crenças razoáveis. A questão chave neste sentido é o tratamento da incerteza. A teoria da decisão normativa ortodoxa, a *teoria da utilidade esperada (UE)*, diz essencialmente que, em situações de incerteza, deve-se preferir a opção com maior desejabilidade ou valor **esperado** (observe que, neste contexto, “desejabilidade” e “valor” devem ser entendidos como desejabilidade / valor de **acordo com o agente em questão**). Esta simples máxima será o foco de grande parte da nossa discussão.

A estrutura deste verbete é a seguinte: A seção 1 discute a noção básica de “preferências sobre prospectos”, que está no âmago da teoria da decisão. A seção 2 descreve o desenvolvimento da teoria da decisão normativa em termos de medidas de preferências cada vez mais poderosas e flexíveis. A seção 3 discute as duas versões mais conhecidas da teoria da UE. A seção 4 considera o significado mais amplo da teoria da UE para a ação prática, inferência e avaliação. A Seção 5 volta-se para os desafios proeminentes da teoria da UE, enquanto a Seção 6 aborda decisões sequenciais e como este cenário mais rico se enquadra nos debates sobre preferências racionais.

1. O que são preferências sobre prospectos?

Os dois conceitos centrais na teoria da decisão são preferências e prospectos (ou o equivalente, opções)⁵. Grosso modo, dizemos (neste verbete) que um agente “prefere” a “opção” A em relação à B no caso de, para o agente em questão, A ser mais desejável ou digna de escolha do que B. Esta definição aproximada deixa

⁵ N.T.: Embora, após a tradução do livro de divulgação científica *Rápido e Devagar*, de Daniel Kahneman, a tradução do termo prospect por “perspectiva” tenha se disseminado, “prospecto” ainda é o que melhor guarda o sentido com que prospect é usado na literatura em economia, que se refere a uma noção mais técnica de “opção” - sem a ambiguidade que a expressão “perspectiva” tem na linguagem comum, na qual denota outras noções, como “ponto de vista”.

claro que a preferência é uma atitude comparativa, a saber, a atitude de comparar opções em termos de quão desejáveis elas são. Além disso, há espaço para discussão sobre o que são as preferências em relação às opções, ou em outras palavras, o que nos interessa quando falamos das preferências de um agente (ou de nós mesmos) em relação às suas opções. Esta seção considera algumas questões elementares de interpretação que preparam o terreno para a introdução (na próxima seção) das tabelas de decisão e da regra de utilidade esperada que, para muitos, é o tópico familiar da teoria da decisão. Outras questões interpretativas relativas a preferências e prospectos serão abordadas posteriormente, à medida que surgirem.

Procedamos, no entanto, primeiro introduzindo candidatas a propriedades básicas da preferência (racional) sobre as opções, para só então voltar às questões de interpretação. Como já mencionado, a preferência diz respeito à comparação entre opções; ela é uma relação entre opções. Para um domínio de opções, falamos de ordenação de preferências de um agente, sendo ela a ordem das alternativas que é gerada pela preferência do agente entre quaisquer duas opções naquele domínio.

No que se segue, $\leq$ representa uma relação de preferência fraca, ou seja, a relação "... não é preferida a ...". Logo, $A \leq B$ significa que o agente em que estamos interessados considera a opção B ao menos tão preferível quanto a A. A partir da preferência fraca se pode definir a preferência estrita, $<$, como segue: $A < B \Leftrightarrow A \leq B \text{ \& } \neg (B \leq A)$ – onde " $\neg X$ " significa "não é o caso que X". Já a relação de indiferença $\sim$ é definida como: $A \sim B \Leftrightarrow A \leq B \text{ \& } B \leq A$ ⁶. Isto significa que o agente que nos interessa está igualmente interessado em A e B.

Dizemos que $\leq$ ordena fracamente um conjunto S de opções quando quer que satisfaça os seguintes axiomas:

Axioma 1 (Completeness)

Para qualquer $A, B \in S$: ou $A \leq B$, ou $B \leq A$.

Axioma 2 (Transitividade)

Para qualquer $A, B, C \in S$: se $A \leq B$ e $B \leq C$, então $A \leq C$

⁶ N.T.: Ou seja, **preferir estritamente** A a B implica **preferir fracamente** A a B, mas não o inverso; e **indiferença** entre A e B implica **preferir fracamente** A a B, e, também, o **inverso**.

O exposto pode ser tomado como uma caracterização preliminar da preferência racional sobre as opções. No entanto, mesmo esta caracterização limitada é controversa, e aponta para interpretações divergentes de "preferências sobre prospectos/opções".

Começemos com o axioma Completeness, que diz que um agente pode comparar, em termos da relação de preferência fraca, todos os pares de opções em S . Se a Completeness é ou não uma restrição plausível de racionalidade depende tanto do tipo de opções que estão sendo consideradas, quanto de como interpretamos as preferências sobre essas opções. Se o conjunto de opções inclui todos os tipos de estados de coisas, então a Completeness não é imediatamente convincente. Por exemplo, é questionável se um agente deve ser capaz de comparar a opção pela qual duas pessoas adicionais no mundo são alfabetizadas com a opção pela qual duas pessoas adicionais atingem a idade de sessenta anos. Se, por outro lado, todas as opções do conjunto são bastante semelhantes entre si, digamos, todas as opções são carteiras de investimento, então a completeness é mais convincente. Mas mesmo se nós não restringirmos os tipos de opções em consideração, a questão de saber se a Completeness deve ou não ser satisfeita depende do significado de preferência. Por exemplo, se as preferências representam apenas comportamento de escolha ou disposições de escolha, como fazem em concordância com a "teoria da preferência revelada" popular entre os economistas (*vide* SEN, 1973), então a completeness é automaticamente satisfeita, assumindo que uma escolha deve ser inevitavelmente feita. Em contraste, se as preferências são entendidas mais como atitudes mentais, ou seja, elas são tipicamente consideradas juízos sobre se uma opção é melhor ou mais desejável do que outra, então as dúvidas sobre a Completeness aludidas acima são pertinentes. Para uma discussão mais aprofundada, consulte Mandler (2001).

A maioria dos filósofos e teóricos da decisão subscrevem esta última interpretação de preferência como uma espécie de julgamento que explica, em vez de ser idêntica a algo, as disposições de escolha e o comportamento de escolha resultante (*vide* HAUSMAN 2011a, 2011b; DIETRICH; LIST, 2016a, 2016b; BRADLEY, 2017. Consulte Thoma (2020b) e Vredenburg (2020) para defesas recentes da **teoria da preferência revelada**, pelo menos no contexto da economia empírica. Além disso, muitos sustentam que a completeness não é racionalmente exigida, tendo em vista que a racionalidade exige apenas os julgamentos que um agente realmente realiza, mas nada diz sobre se um julgamento deve ser realizado em primeiro lugar. No entanto, seguindo Richard Jeffrey (1983), a maioria dos teóricos das decisões

sugere que a racionalidade exige que as preferências sejam **coerentemente extensíveis**. Isso significa que, mesmo que suas preferências não sejam completas, deve ser possível completá-las sem violar nenhuma das condições racionalmente exigidas, em particular, a Transitividade.

Isto nos leva ao axioma da Transitividade, que diz que se uma opção B é fracamente preferível a A, e C é fracamente preferível a B, então C é fracamente preferível a A. Um desafio recente à Transitividade gira em torno de conjuntos heterogêneos de opções, conforme a discussão da Completude acima. Mas aqui se faz uma interpretação diferente de preferência na comparação de opções. A ideia é que as preferências, ou julgamentos de desejabilidade, podem ser sensíveis a uma condição de saliência. Por exemplo, suponha que a característica mais saliente ao comparar os carros A e B seja o quão rápido eles podem ser dirigidos, e B não é pior do que A neste aspecto; ainda assim, a característica mais saliente ao comparar carros B e C é o quão seguros eles são, e C não é pior do que B neste aspecto. Além disso, quando se comparam A e C, a característica mais marcante é quão belos eles são. Neste caso, alguns argumentam (TEMKIN, 2012) que não há razão para que a Transitividade deva ser satisfeita com respeito às preferências relativas a A, B e C. Outros (BROOME, 1991a) argumentam que a Transitividade é parte do próprio significado da relação “melhor que” (ou desejabilidade comparativa objetiva); se a preferência racional é um julgamento de melhoria ou desejabilidade, então a Transitividade não é negociável. Com relação ao exemplo do carro, Broome argumentaria que a desejabilidade de uma opção totalmente especificada não deve variar, simplesmente em virtude das outras opções com as quais ela é comparada. Isso quer dizer que ou o contexto de escolha afeta como o agente percebe a opção em questão, caso em que a descrição da opção deve refletir isso, ou então o contexto de escolha não afeta a opção. De qualquer forma, a Transitividade deve ser satisfeita.

Há uma defesa mais direta da Transitividade de preferências; uma defesa que depende das perdas certas que podem ocorrer a qualquer um que viole o axioma. Este é o chamado argumento **da bomba de dinheiro** (*money pump*) – confira Davidson *et. al.* (1955) para um argumento inicial deste tipo; mas, para discussão e revisão recentes deste argumento, consulte Gustafsson (2010; 2013). Ele se baseia na suposição de que se você acha X pelo menos tão desejável quanto Y, então você deve ficar feliz em trocar o último pelo primeiro. Suponha que você viole a Transitividade, ou seja, por você, pois, para você: $A \leq B$, $B \leq C$, mas $C < A$. Além disso, suponha que você tenha atualmente a opção A. Então você deve estar disposto a trocar A por B. O mesmo vale para B e C: você deve estar disposto a

trocar B por C. Você prefere estritamente A a C, então você deve estar disposto a trocar C, mais alguma coisa (p. ex., alguma, quantia de dinheiro), por A. Mas agora você está na mesma situação em que você começou, a saber, tendo A, e não tendo B nem C – exceto pelo fato de que você ficou mais pobre! Assim, em poucos passos, cada um deles coerente com suas preferências, você se encontra em uma situação claramente pior, a partir de seu próprio ponto de vista, do que a situação original. O quadro se torna mais dramático se imaginarmos que o processo poderia ser repetido, transformando você em uma "bomba de dinheiro". Portanto, o argumento segue, há algo (instrumentalmente) irracional em suas preferências intransitivas. Se suas preferências fossem transitivas, então você não estaria vulnerável a escolher uma opção **dominada** e a servir como uma bomba de dinheiro. Portanto, suas preferências devem ser transitivas.

Embora as controvérsias acima mencionadas não tenham sido resolvidas, as seguintes hipóteses serão feitas no restante deste artigo: i) os objetos de preferência podem ser prospectos heterogêneos, incorporando um domínio rico e variado de propriedades, ii) a preferência entre opções é um julgamento de conveniência comparativa ou de capacidade de escolha, e iii) as preferências satisfazem tanto a Completude quanto a Transitividade (embora esta última condição seja revisitada na Seção 5). A questão que agora se coloca é se existem outras restrições gerais à preferência racional sobre as opções.

2. Medidas de utilidade para as preferências

Continuando nossa investigação das preferências racionais sobre os prospectos, a **representação** numérica (ou **medida**) das ordenações de preferência se tornará importante. As medidas numéricas em questão são conhecidas como **funções de utilidade**. Os dois principais tipos de função de utilidade que desempenharão um papel são a função de utilidade **ordinal** e a função de utilidade mais rica em informações, **com valores em intervalos** (ou **cardinal**).

2.1 Utilidade ordinal

Acontece que, uma vez que o conjunto de prospectos/opções S é finito, qualquer ordenação fraca das opções em S pode ser representada por uma função

de utilidade ordinal. Para ser mais preciso, digamos que u é uma função de utilidade de domínio S . Dizemos que a função u representa a preferência $\leq$ entre as opções em S somente se:

1. Para qualquer $A, B \in S$: $u(A) \leq u(B) \Leftrightarrow A \leq B$

Dito de outro modo, quando o acima exposto se mantém, a relação de preferência pode ser representada como **maximização da utilidade**, uma vez que ela sempre favorece a opção com maior utilidade.

A única informação contida em uma representação de utilidade ordinal é como o agente cujas preferências estão sendo representadas ordena as opções, a saber, da menos para a mais preferível. Isso significa que se u é uma função de utilidade ordinal que representa a ordenação $\leq$, então qualquer função de utilidade u' que seja uma transformação ordinal de u – isto é, qualquer transformação que também satisfaça o bicondicional em (1) – representa $\leq$ igualmente bem. Portanto, dizemos que uma função de utilidade ordinal é **única apenas até às transformações ordinais**.

O resultado apresentado acima pode ser resumido da seguinte forma:

Teorema 1 (Representação Ordinal). Seja S um conjunto finito e $\leq$ uma relação de preferência fraca em S . Então, há uma função de utilidade ordinal que representa $\leq$ apenas quando $\leq$ é completa e transitiva.

Este teorema não deve ser muito surpreendente. Se $\leq$ é completa e transitiva com relação a S , então as opções em S podem ser colocadas em uma ordem, da mais para a menos preferida, em que algumas opções podem cair na mesma posição (se forem consideradas igualmente desejáveis), mas sem haver ciclos, *loops* ou lacunas. O teorema 1 diz apenas que podemos atribuir números às opções de uma forma que represente esta ordem. Para uma prova simples do Teorema 1, exceto por uma relação de preferência rígida em vez de fraca, consulte Peterson (2009, p. 95).

Note que as utilidades ordinais não são, por assim dizer, matematicamente muito "poderosas". Não faz sentido, por exemplo, comparar as expectativas probabilísticas de diferentes conjuntos de utilidades ordinais. Por exemplo, considere os seguintes dois pares de prospectos: aos elementos do primeiro par são atribuídas

utilidades ordinais de 2 e 4, enquanto aos do segundo par são atribuídas utilidades ordinais de 0 e 5. Vamos especificar uma distribuição de probabilidade "plana" em cada caso, de modo que cada elemento dos dois pares corresponda a uma probabilidade de 0.5. Com relação a esta distribuição de probabilidade, a expectativa do primeiro par de utilidades ordinais é 3, que é maior que a expectativa de 2.5 do segundo par. No entanto, quando transformamos as utilidades ordinais de uma forma admissível – por exemplo, aumentando a utilidade mais alta do segundo par de 5 para 10 – a ordenação das expectativas se inverte; agora a comparação é entre 3 e 5. O significado deste ponto ficará mais claro a seguir, quando nos voltarmos à avaliação comparativa das loterias e escolhas arriscadas. Uma função de utilidade intervalada ou cardinal é necessária para avaliar loterias/prospectos de risco entre si de uma forma consistente. Da mesma forma, a fim de construir ou conceituar uma função de utilidade cardinal, tipicamente apela-se para preferências sobre as loterias. A esse respeito, consulte Alt (1936) para uma construção "sem risco" de utilidade cardinal, ou seja, uma que não apela à loterias.

2.2 Utilidade Cardinal

Para obter uma representação cardinal (com valores em intervalos) de utilidade para uma ordenação de preferências – ou seja, uma medida que representa não apenas a forma como um agente ordena as opções, mas também diz algo sobre a "distância" do desejo entre as opções – precisamos de uma configuração mais rica; o conjunto de opções e a correspondente ordem de preferência precisará ter mais estrutura do que a de uma medida de utilidade ordinal. Uma dessas abordagens, devida a John von Neumann e Oskar Morgenstern (1944), será exposta em detalhe abaixo. Por enquanto, é útil focar no tipo de opção que é fundamental para entender e construir uma função utilitária cardinal: as loterias⁷.

Considere primeiro uma ordenação entre três opções comuns, por exemplo, três destinos de férias: Amsterdã, Bangkok e Cardiff, indicados respectivamente por A, B e C. Suponha que a ordem de suas preferências seja $A < B < C$. Esta informação é suficiente para representar ordinalmente seu juízo; lembre-se que

⁷ Nossa discussão sobre a cardinalização de utilidades é bastante semelhante à proposta por Resnik (1987, p. 88-91).

qualquer atribuição de utilidades é aceitável, desde que C receba um valor maior que B, e B um valor maior que A. Mas talvez queiramos saber mais do que o que se pode inferir de tal função de utilidade – queremos saber o quanto C é preferido em relação a B, comparado com o quanto B é preferido em relação a A. Por exemplo, pode ser que Bangkok seja considerada quase tão desejável quanto Cardiff, mas Amsterdã esteja muito longe de Bangkok, relativamente falando. Ou talvez Bangkok seja apenas marginalmente melhor do que Amsterdã, em comparação com a medida em que Cardiff é melhor que Bangkok. Este tipo de informação sobre a distância relativa entre opções, em termos de força de preferência ou desejabilidade, é precisamente o que é dado por uma função de utilidade com valores em intervalos (ou intervalados). O problema é como estabelecer esta informação.

Para resolver este problema, Ramsey (1926) e, mais tarde, von Neumann e Morgenstern (doravante vNM) fizeram a seguinte sugestão: construímos uma nova opção, uma **loteria L** que tem A e C como possíveis "prêmios", e descobrimos quais as chances que a loteria deve dar para que você fique indiferente entre esta loteria e passar as férias em Bangkok. A ideia básica é que o seu julgamento sobre Bangkok, em relação a Cardiff, por um lado, e Amsterdã, por outro, pode ser medido pelo nível de risco que a loteria deve distribuir entre Cardiff e Amsterdã até você a considerar tão desejável quanto Bangkok. Por exemplo, se você é indiferente entre Bangkok e uma loteria que oferece uma chance muito baixa de ganhar uma viagem a Cardiff, então você evidentemente não considera Bangkok muito melhor que Amsterdã, em relação a Cardiff; para você, mesmo uma pequena melhoria em relação a Amsterdã, ou seja, uma loteria com uma pequena chance de Cardiff em vez de Amsterdã, é suficiente para se igualar a Bangkok.

A análise acima pressupõe que as loterias são avaliadas em termos de sua capacidade de escolha ou desejabilidade **esperada**. Ou seja, a desejabilidade de uma loteria é, de fato, a soma das chances de cada prêmio multiplicada pela desejabilidade desse prêmio. Considere o seguinte exemplo: suponha que você fique indiferente entre a loteria e o feriado em Bangkok quando a chance da loteria resultar em um feriado em Cardiff é de $3/4$. Chamemos esta loteria peculiar de **L'**. A ideia é que Bangkok está, portanto, a três quartos da escala de desejabilidade que tem Amsterdã na parte inferior e Cardiff na parte superior. Se estipularmos que $u(A)=0$ e $u(C)=1$, então $u(B) = u(L') = 3/4$. Isto corresponde à desejabilidade **esperada** – ou, como é normalmente chamado, a **utilidade esperada** – da loteria, uma vez que $u(L') = 1/4 \cdot 0 + 3/4 \cdot 1 = 3/4$. Ou seja, o valor da loteria é uma soma ponderada de probabilidade das utilidades de seus prêmios, onde o peso de cada prêmio é

determinado pela probabilidade de que a loteria resulte nesse prêmio.

Assim, vemos que uma medida de utilidade com valores em intervalos sobre as opções pode ser construída através da introdução de opções de loteria. Como o nome sugere, a medida da utilidade em valores em intervalos transmite informações sobre os tamanhos relativos dos intervalos entre as opções de acordo com alguma escala de desejabilidade. Ou seja, as utilidades são únicas após termos fixado o ponto de partida da nossa medida e a unidade da escala de desejabilidade. No exemplo acima, poderíamos, por exemplo, ter atribuído um valor de utilidade de 1 a A, e 5 a C e neste caso teríamos que atribuir um valor de utilidade de 4 a B, já que 4 é $3/4$ do caminho entre 1 e 5. Em outras palavras, uma vez atribuído o valor da utilidade a A e C, a utilidade de L' e, conseqüentemente, de B estão assim determinadas. Vamos chamar esta segunda função de utilidade u' . Ela está relacionada à nossa função original como segue: $u' = 4 \cdot u + 1$. Esta relação sempre se mantém entre duas funções desse tipo: se u é uma função de utilidade com valores em intervalo que representa a ordenação de preferências, $\leq$, e u' é uma outra função de utilidade que também representa esta mesma ordenação de preferência, então existem constantes a e b , em que a deve ser positivo, tal que $u' = a \cdot u + b$. Isto equivale a dizer que funções de utilidade com valor de intervalo são únicas apenas até a transformação linear positiva.

Antes de concluir esta discussão sobre medição da utilidade, devem ser mencionadas duas limitações relacionadas com as informações que tais medidas transmitem. Em primeiro lugar, como as utilidades das opções, sejam elas ordinais ou valoradas em intervalos, só podem ser determinadas **em relação** às utilidades de outras opções; não existe a utilidade **absoluta** de uma opção, pelo menos não sem outras suposições⁸. Segundo, pelo mesmo raciocínio, nem as medidas de utilidade cardinais, nem as medidas de utilidade ordinárias, como aqui discutidas, são **interpessoalmente comensuráveis** no que diz respeito aos níveis e unidades de utilidade. A título de ilustração rápida, suponha que tanto você quanto eu tenhamos a ordenação de preferência descrita acima sobre as opções de férias: $A < B < C$. Suponha também que, de acordo com o acima descrito, nós dois somos indiferentes

⁸ As premissas adicionais precisariam relacionar opções particulares a níveis privilegiados particulares de utilidade; por exemplo, seria necessário argumentar que a ordenação de preferência de um agente racional deve incorporar, digamos, uma opção privilegiada de utilidade zero, caso em que as proporções de distâncias de utilidade dessa opção seriam significativas.

entre **B** e a loteria **L'** que tem $3/4$ de chance de resultar em **C** e $1/4$ de chance de resultar em **A**. Podemos então dizer que me dar a opção de Cardiff, e a você a de Bangkok, equivaleria à mesma quantidade de "desejabilidade total" que conceder a você Cardiff e a mim Bangkok? Nós não temos condições de concluir isto. Nossa preferência compartilhada é, por exemplo, coerente com o fato de eu achar que passar as férias em Cardiff é um sonho tornado realidade enquanto você simplesmente acha isso o melhor de um lote de opções ruins. Além disso, nem temos o direito de dizer que a diferença de desejo entre Bangkok e Amsterdã é a mesma para você e para mim. Da minha perspectiva, o desejo das três opções pode variar desde viver o inferno até um sonho se tornando realidade, enquanto da sua, varia de ruim a muito ruim; ambas as avaliações são consistentes com a ordenação de preferências acima. Na verdade, ele pode ser válido para nossas preferências sobre **todas** as opções possíveis, incluindo as loterias: mesmo que compartilhemos a mesma ordenação total de preferências, pode ser que você esteja apenas com uma disposição negativa (não encontrando nenhuma opção tão boa), enquanto eu sou muito extremista – encontrando algumas opções excelentes, mas outras uma verdadeira tortura. Assim, as funções de utilidade, sejam elas de valor intervalado ou ordinais, não permitem comparações interpessoais significativas. No trabalho de Elster e Roemer (1993) há uma série de artigos discutindo essas questões. Consulte, também, o verbete da SEP sobre **Teoria da Escolha Social**.

2.3 O teorema da representação de von Neumann e Morgenstern (vNM)

A última seção forneceu uma representação das utilidades, com valores em intervalos, das preferências de uma pessoa em relação às loterias, na suposição de que as loterias são avaliadas em termos de utilidade esperada. Alguns podem achar isso um pouco rápido demais. Por que devemos assumir que as pessoas avaliam as loterias em termos de utilidades esperadas? O teorema vNM efetivamente suprime as lacunas no raciocínio ao chamar a atenção de volta para a relação de preferências. Além da Transitividade e da Completude, o teorema vNM introduz outros princípios que regem as preferências racionais sobre as loterias, e mostra que as preferências de um agente podem ser representadas como maximizando a utilidade esperada sempre que suas preferências satisfazem esses princípios.

Vamos primeiro definir, em termos formais, a utilidade esperada de uma loteria: seja L_i uma loteria do conjunto **L** de loterias, O_{ik} o resultado, ou prêmio, da

loteria L_i que acontece com probabilidade p_{ik} . A utilidade esperada de L_i é então definida como:

A equação vNM

$$EU(L_i) = \sum_k u(O_{ik}) \cdot p_{ik}$$

A suposição feita anteriormente pode agora ser formalmente afirmada:

1. Para qualquer $L_i, L_j \in L$: $L_i \leq L_j \Leftrightarrow EU(L_i) \leq EU(L_j)$

Quando o acima exposto se aplica, dizemos que existe uma função de utilidade esperada que representa as preferências do agente; em outras palavras, o agente pode ser representado como **maximizando a utilidade esperada**. A questão que vNM aborda é: Que tipos de preferências podem ser assim representados? Para responder a esta pergunta, devemos voltar à relação de preferência subjacente ($\leq$) sobre o conjunto de opções; neste caso envolvendo loterias. O teorema vNM exige que o conjunto de loterias seja bastante extenso: ele é "fechado sob mistura de probabilidade", isto é, se $L_i, L_j \in L$ então loterias compostas que tenham L_i e L_j como possíveis prêmios também pertencem a L (outro pressuposto técnico, que não será discutido em detalhe, é que as loterias compostas podem sempre ser reduzidas, de acordo com as leis de probabilidade, a simples loterias que envolvam apenas prêmios básicos).

Uma restrição básica de racionalidade na relação de preferência já foi discutida - ela ordena fracamente as opções (ou seja, ela satisfaz a transitividade e a completude). A seguinte notação será utilizada para introduzir os dois axiomas vNM adicionais: $\{pA, (1-p)B\}$ denota uma loteria que resulta ou em A, com probabilidade p , ou B, com probabilidade $1-p$.

Axioma 3 (Continuidade)

Suponha $A \leq B \leq C$. Logo existe $p \in [0, 1]$ tal que:
 $\{pA, (1-p)C\} \sim B$

Axioma 4 (Independência)

Suponha $A \leq B$. Logo, para qualquer C e qualquer $p \in [0, 1]$:
 $\{pA, (1-p)C\} \leq \{pB, (1-p)C\}$

Continuidade implica que nenhum resultado A é tão ruim que você não estaria disposto a fazer alguma aposta que poderia resultar em você terminar com esse resultado, mas poderia resultar ao invés disso em você terminar com um resultado (C) que você acha ser uma melhoria marginal em relação a seu *status quo* (B), desde que a chance de A seja pequena o suficiente. Intuitivamente, a Continuidade garante que as avaliações de um agente de loterias sejam adequadamente sensíveis às probabilidades dos prêmios das loterias.

A Independência implica que, quando duas alternativas têm a mesma probabilidade para algum resultado em particular, nossa avaliação das duas alternativas deve ser independente da nossa opinião sobre esse resultado em particular. Intuitivamente, isto significa que as preferências entre loterias devem ser regidas apenas pelas características em que as loterias diferem; as semelhanças entre as loterias devem ser efetivamente ignoradas.

Algumas pessoas consideram o axioma da Continuidade uma restrição injustificada à preferência racional. Existe alguma probabilidade p tal que você estaria disposto a aceitar uma aposta com a probabilidade $(1 - p)$ de você ganhar \$10 e a probabilidade de você perder sua vida? Muitas pessoas pensam que não há. No entanto, as mesmas pessoas presumivelmente atravessariam a rua para pegar uma nota de 10 dólares que eles haviam deixado cair. Porém, isso é apenas uma aposta que tem uma probabilidade muito pequena de morte por atropelamento, mas uma probabilidade muito maior de ganhar \$10! De modo mais geral, embora as pessoas raramente pensem assim, elas constantemente fazem apostas que têm chances minúsculas de levar à morte iminente e chances correspondentes muito altas de alguma recompensa modesta. Toda vez que damos uma volta, dirigimos nosso carro, vamos a algum lugar, e assim por diante, há alguma chance de termos um acidente fatal. Mas, como as probabilidades desses acidentes são suficientemente pequenas, decidimos correr riscos.

A independência parece uma exigência de racionalidade quando considerada em abstrato. No entanto, há exemplos famosos onde as pessoas frequentemente violam a Independência sem parecerem irracionais. Esses exemplos envolvem **complementaridades** entre os possíveis resultados da loteria. Um exemplo particularmente conhecido é o chamado *Paradoxo de Allais* que o economista francês Maurice Allais (1953) introduziu pela primeira vez no início da década de 1950. O paradoxo gira em torno da comparação das preferências das pessoas sobre dois pares de loterias similares aos dados na Tabela 1. As loterias são descritas em termos dos prêmios associados a determinados bilhetes numerados, onde um

bilhete será sorteado aleatoriamente (por exemplo, resulta em um prêmio de US\$2500 se um dos bilhetes numerados de 2 a 34 for sorteado).

	1	2-34	35-100
L1	\$0	\$2500	\$2400
L2	\$2400	\$2400	\$2400

	1	2-34	35-100
L3	\$0	\$2500	\$0
L4	\$2400	\$2400	\$0

Tabela 1. O paradoxo de Allais

Nessa situação, muitas pessoas preferem estritamente L2 a L1, mas também L3 a L4 (como evidenciado pelo seu comportamento de escolha e pelo seu testemunho), um par de preferências que serão chamadas de preferências de Allais⁹. Uma maneira comum de racionalizar as preferências de Allais é que, na situação da primeira escolha, o risco de acabar sem nada ao invés de ter \$ 2400 garantidos não justifica o aumento das chances de um prêmio maior. Na situação da segunda escolha, no entanto, o mínimo que se pode ganhar é \$0, não importa qual escolha se faça. Portanto, nesse caso muitas pessoas pensam que o leve risco extra de \$0 vale a chance de um prêmio melhor.

Embora o raciocínio acima possa parecer convincente, as preferências de Allais entram em conflito com o axioma da Independência. O seguinte é verdadeiro para ambas as situações de escolha: qualquer que seja a sua escolha, você receberá o mesmo prêmio se um dos ingressos da última coluna for sorteado. Portanto, a Independência implica que tanto a sua preferência entre L1 e L2 e sua preferência entre L3 e L4 devem ser independentes dos prêmios nessa coluna. Mas quando você ignora a última coluna, L1 se torna idêntico a L3 e L2 a L4. Portanto, se você preferir L2 a L1, mas L3 a L4, parece haver uma inconsistência na sua ordenação de preferência. E, definitivamente, há uma violação da Independência. Como

⁹ Veja Heap *et al.* (1992, cap. 3) para uma discussão acessível dos resultados empíricos sobre o paradoxo de Allais, bem como alguns modelos empíricos sugeridos em resposta ao paradoxo.

resultado, o par de preferências em discussão não pode ser representado como a maximização da utilidade esperada (por isso, o "paradoxo": muitas pessoas pensam que a Independência é uma exigência de racionalidade, mas querem também afirmar que não há nada de irracional nas preferências de Allais).

Os teóricos da decisão têm reagido de diferentes maneiras ao paradoxo de Allais. Essa questão será revisitada na Seção 5.2, quando os desafios à teoria da UE forem discutidos. O presente objetivo é simplesmente mostrar que Continuidade e Independência são restrições convincentes à preferência racional, embora elas tenham seus detratores. O resultado vNM provado pode ser resumido assim:

Teorema 2 (von Neumann-Morgenstern – vNM)

Seja $\mathbf{O}$ um conjunto finito de resultados, $\mathbf{L}$ um conjunto de loterias correspondentes que é fechado sob mistura de probabilidade e $\leq$ uma relação de preferência fraca em $\mathbf{L}$. Então, $\leq$ satisfaz os axiomas 1-4 se, e somente se, existir uma função u de $\mathbf{O}$ no conjunto dos números reais, que é única até transformação linear positiva, e em relação à qual $\leq$ pode ser representada como maximizando a utilidade esperada.

David Kreps (1988) fornece uma ilustração acessível da prova deste teorema.

3. Tomando decisões reais

O teorema vNM é um resultado muito importante para medir a força das preferências de um agente racional sobre opções seguras (as loterias efetivamente facilitam representar uma medida cardinal sobre opções seguras). Mas isto não nos leva a tomar decisões racionais no mundo real; ainda não temos uma teoria de decisão. O teorema limita-se a avaliar opções que vêm com uma distribuição objetiva de probabilidade sobre resultados – uma situação que os teóricos e economistas muitas vezes descrevem como “escolha sob risco” (KNIGHT, 1921).

Na maioria das situações de escolha comum, os objetos de escolha, sobre os quais devemos ter ou formar preferências, não são assim. Ao contrário, os tomadores de decisão devem consultar **suas próprias** crenças sobre a probabilidade de um ou outro desfecho resultar de uma determinada opção. Decisões em tais

circunstâncias são frequentemente descritas como “escolhas sob incerteza” (KNIGHT, 1921). Por exemplo, considere a situação de uma montanhista decidindo se vai ou não tentar uma perigosa subida ao topo, na qual o fator chave para ela é o clima. Se ela tiver sorte, poderá ter acesso às estatísticas meteorológicas abrangentes para a região. No entanto, as estatísticas meteorológicas diferem da configuração da loteria, pois não **determinam** as probabilidades dos possíveis resultados de tentar *versus* não tentar alcançar o cume em um determinado dia. Não menos importante, a montanhista deve considerar quão confiante está no procedimento de coleta de dados, se as estatísticas são aplicáveis ao dia em questão, e assim por diante, ao avaliar suas opções em função do tempo.

Alguns dos resultados mais celebrados na teoria da decisão abordam, até certo ponto, esses desafios. Eles consistem em mostrar quais condições nas preferências sobre “opções do mundo real” são suficientes para a existência de um par de funções de utilidade e probabilidade em relação ao qual o agente pode ser representado como maximizando a utilidade esperada. A interpretação padrão é que, assim como a função de utilidade representa os desejos do agente, a função de probabilidade representa suas crenças. As teorias são referidas coletivamente como *teoria da utilidade esperada subjetiva (SEU)*, pois dizem respeito às preferências de um agente sobre prospectos que são caracterizadas inteiramente em termos de suas próprias crenças e desejos (mas continuaremos a usar o rótulo mais simples de *teoria da UE*). Nesta seção, dois desses resultados serão brevemente discutidos: o de Leonard Savage (1954) e o de Richard Jeffrey (1965).

Note que estas teorias de decisão da UE aparentemente prescrevem duas coisas: (a) você deve ter atitudes de preferência consistentes, e (b) você deve preferir os meios para seus fins, ou pelo menos você deve preferir os meios que você avalia que levarão, **em média**, aos seus fins (BUCHAK, 2016). A questão que se coloca é: Qual a relação entre essas receitas? Os *teoremas de representação da UE* que serão delineados em breve parecem mostrar que, apesar das aparências, as duas prescrições são na verdade apenas uma: qualquer pessoa que tenha atitudes consistentes prefere os meios aos seus fins, e vice-versa. Mas o enigma continua sendo que há muitas maneiras de ter atitudes de preferência consistentes, e certamente nem todas elas equivalem a preferir os meios aos seus **próprios** fins **verdadeiros**. Vale a pena ter em mente este quebra-cabeças ao avaliar a teoria da UE em suas diversas facetas; ele surgirá novamente mais tarde.

3.1 A teoria de Savage

A teoria da decisão de Leonard Savage, como apresentada em seu (1954) *The Foundations of Statistics*, é sem dúvida a mais conhecida teoria normativa de escolha sob incerteza; em particular dentro da economia e das ciências da decisão. No livro, Savage apresenta um conjunto de axiomas que restringem as preferências sobre um conjunto de opções, garantindo a existência de um par de funções de probabilidade e utilidade em relação ao qual as preferências podem ser representadas como maximização da utilidade esperada. Quase três décadas antes da publicação do livro, Frank P. Ramsey (1926) havia proposto que um conjunto diferente de axiomas poderia gerar mais ou menos o mesmo resultado. No entanto, a teoria de Savage tem sido muito mais influente que a de Ramsey, talvez porque Ramsey não deu uma prova completa de seu resultado, nem forneceu muitos detalhes de como esta seria (BRADLEY, 2004). O resultado de Savage não será descrito aqui em todos os seus detalhes. Entretanto, os ingredientes e a estrutura de seu teorema serão expostos destacando seus pontos fortes e fracos.

As opções ou prospectos na teoria de Savage são semelhantes às loterias, exceto que os possíveis resultados não vêm com chances objetivas, mas dependem de se um determinado estado de coisas do mundo é real. De fato, os elementos primitivos na teoria de Savage são **resultados**¹⁰ e **estados (do mundo)**. Os primeiros são os bons ou maus estados de coisas que, em última análise, afetam e importam para um agente, enquanto os segundos são as características do mundo sobre as quais o agente não tem controle e que são o *locus* de sua incerteza sobre o mundo. Conjuntos de estados são chamados de **eventos**. Esta distinção entre resultados e estados serve para separar claramente desejo e crença: os primeiros são, de acordo com a teoria de Savage, o alvo do desejo, enquanto os segundos são o alvo da crença.

As opções semelhantes a loterias sobre as quais o agente tem preferências são um rico conjunto de *atos* que efetivamente equivalem a todas as possíveis atribuições de resultados aos estados do mundo. Ou seja, os “atos” são funções do **espaço de estados** para o **espaço de resultados**, e a ordenação de preferências do agente é considerada como definida sobre todas essas possíveis funções. Alguns desses “atos” parecerão bastante sensatos: considere o ato que atribui ao evento

¹⁰ O próprio Savage usa o termo “consequências” ao invés de “resultados”.

"chove" o resultado "passeio molhado e miserável" e atribui ao evento "não chove" o resultado "passeio muito confortável". Este é aparentemente o ato de dar um passeio sem o guarda-chuva. Outros atos de Savage não parecerão tão sensatos, como o *ato constante* que atribui ao evento "chove" e "não chove" o mesmo resultado "passeio molhado e miserável" (note que os atos constantes proporcionam uma forma de incluir resultados seguros dentro da ordenação de preferência). O problema com este ato (e muitos outros) é que ele não corresponde a nada que um agente poderia, sequer em princípio, optar por fazer ou executar¹¹.

A distinção de Savage entre ato/estado(evento)/resultado pode ser naturalmente representada em forma de tabela, com linhas servindo como atos que produzem um determinado resultado para cada coluna de estado/evento. A Tabela 2 mostra os dois atos mencionados acima, mais um terceiro com que o tomador de decisão pode se importar: os atos i) "passear sem guarda-chuva", ii) "passear com guarda-chuva", e iii) o bizarro "ato constante". É claro que o conjunto de atos necessários para o teorema de Savage envolve ainda mais atos que respondem por todas as combinações possíveis de estados e resultados.

	Sem chuva	Chuva
Passear sem guarda-chuva	Passeio muito confortável	Passeio molhado e miserável
Passear com guarda-chuva	Passeio confortável	Passeio confortável
Ato constante	Passeio molhado e miserável	Passeio molhado e miserável

Tabela 2. Tabela de decisão ao estilo de Savage

Antes de discutir os axiomas de Savage, falemos do resultado a que eles dão origem. A seguinte notação será utilizada: f , g , etc, são vários atos, ou seja, funções do conjunto $\mathbf{S}$ de estados do mundo no conjunto $\mathbf{O}$ de resultados, sendo o conjunto destas funções $\mathbf{F}$. $f(s_j)$ denota o resultado de f quando o estado $s_j \in \mathbf{S}$ é verdadeiro. A utilidade esperada de f , de acordo com a teoria de Savage, designada por $U(f)$, é dada pela

¹¹ Este problema de atos sem sentido é considerado por alguns o principal problema com o resultado de Savage (vide JOYCE, 1999, p. 108). Levantamos essa questão novamente abaixo quando discutimos os pontos fracos da teoria de Savage.

Equação de Savage

$$U(f) = \sum_i u(f(s_i)) \cdot P(s_i)$$

Savage demonstrou que o resultado pode ser enunciado da seguinte forma¹²:

Teorema 3 (Savage).

Seja $\leq$ uma relação de preferência fraca em F . Se $\leq$ satisfizer os axiomas de Savage, então o seguinte se aplica:

- A confiança do agente na atualidade dos estados em S pode ser representada por uma função de probabilidade única (e finitamente aditiva) P ;
- a “força” de seus desejos pelos resultados em O pode ser representada por uma função de utilidade, que é única até a transformação linear positiva;
- e o par (P, u) dá lugar à função de utilidade esperada, U , que representa as suas preferências pelas alternativas em F ou seja, para qualquer $f, g \in F$:
 $f \leq g \Leftrightarrow U(f) \leq U(g)$

O resultado acima pode parecer notável; em particular, o fato de que as preferências de uma pessoa podem determinar uma função de probabilidade única que representa suas crenças. Em um olhar mais atento, entretanto, é evidente que algumas de nossas crenças podem ser determinadas pelo exame de nossas preferências. Suponha que lhe seja oferecida uma escolha entre duas loterias, uma que resulte em você ganhar um belo prêmio se uma moeda der cara, mas sem receber nada se a moeda der coroa, e outra que resulte em você ganhar o mesmo prêmio se a moeda der coroa, mas não receber nada se a moeda der cara. Então, assumindo que a desejabilidade do prêmio (e, de modo similar, a de não haver prêmio) é independente de como a moeda cai, sua preferência entre as duas loterias deve ser inteiramente determinada por suas crenças comparativas para as duas maneiras como a moeda pode cair. Por exemplo, se você prefere estritamente a primeira loteria à segunda, isso sugere que você considera cara mais provável do que coroa.

¹² Para simplificar, assumiremos que o conjunto O é finito, mas Savage provou um resultado semelhante para um conjunto O infinito.

A observação acima sugere que se podem medir as crenças comparativas de um agente, e talvez mais, a partir de suas preferências. Savage foi um passo além disso e *definiu* crenças comparativas em termos de preferências. Para afirmar a definição de Savage, seja $\preceq$ uma relação de crença comparativa fraca, definida sobre o conjunto **S** de estados do mundo. ($\prec$ e $\sim$ são definidos em termos de $\preceq$ da forma usual).

Definição 1 (Crença Comparativa).

Suponha que E e F sejam dois eventos (ou seja, subconjuntos de **S**). Suponha que X e Y sejam dois resultados, e f e g dois atos, com as seguintes propriedades:

- $f(s_i)=X$ para todo $s_i \in E$, mas $f(s_i)=Y$ para todo $s_i \notin E$,
- $g(s_i)=X$ para todo $s_i \in F$, mas $g(s_i)=Y$ para todo $s_i \notin F$,
- $Y \preceq X$.

Logo, $E \preceq F \Leftrightarrow f \leq g$.

A definição 1 é baseada na simples observação de que geralmente se prefere associar um bom resultado a um evento mais provável do que a um menos provável. Mas a ideia de que isso **define** crenças comparativas pode parecer questionável. Poderíamos, por exemplo, imaginar pessoas que são instrumentalmente irracionais e, como resultado, deixam de preferir g a f , mesmo quando todas as condições acima se mantêm e elas acham F mais provável do que E . Além disso, esta definição levanta a questão de como definir as crenças comparativas daqueles que são indiferentes entre **todos** os resultados (ERIKSSON; HÁJEK, 2007). Talvez não existam tais pessoas (e o axioma de Savage P5 de fato deixa claro que seu resultado não se refere a tais pessoas). No entanto, parece que uma definição de crenças comparativas não deve impedir que tais pessoas, se existirem, tenham crenças comparativas. Savage sugere que esta definição de crenças comparativas é plausível à luz do seu axioma P4, que será enunciado a seguir. Em qualquer caso, ocorre que, quando as preferências de uma pessoa satisfazem os axiomas de Savage, podemos ler a partir de suas preferências uma relação de crença comparativa que pode ser representada por uma função de probabilidade (única).

Sem mais delongas, vamos enunciar, a seguir, os axiomas de Savage. Estes são entendidos como restrições à relação de preferência de um agente sobre um conjunto de atos, tal como descrito acima. O primeiro dos axiomas de Savage é o axioma básico de ordenação.

P1. (Ordenação)

A relação $\leq$ é completa e transitiva.

O próximo axioma lembra o axioma de Independência de vNM. Dizemos que a alternativa f “concorda com” g se, para qualquer estado no evento E , f e g produzem o mesmo resultado.

P2. Princípio da Coisa Segura (*sure thing principle*)

Se f , g , e f' , g' são tais que:

- f concorda com g e f' concorda com g' no evento $\neg E$,
 - f concorda com f' e g concorda com g' no evento E ,
 - e $f \leq g$,
- então $f' \leq g'$.

A ideia por trás do Princípio da Coisa Segura (PCS) é essencialmente a mesma que está por trás da Independência: como devemos ser capazes de avaliar cada resultado independentemente de outros resultados possíveis, podemos seguramente ignorar estados do mundo onde dois atos que estamos comparando produzem o mesmo resultado. Colocar o princípio em forma de tabela pode tornar isso mais aparente. A configuração envolve quatro atos com a seguinte forma:

	E	$\neg E$
f	X	Z
g	Y	Z
f'	X	W
g'	Y	W

A intuição por trás do PCS é que se g é fracamente preferível a f , então isso deve ser porque a consequência Y é considerada pelo menos tão desejável quanto X ; o que pelo mesmo raciocínio implica que g' é fracamente preferido a f' .

Savage também exige que a conveniência de um resultado seja independente do estado em que ele ocorre. Para formalizar esta exigência, Savage introduz a noção de evento **nulo**, definido da seguinte forma:

Definição 2 (Evento nulo)

O evento E é nulo para qualquer alternativa $f, g \in F$, $f \sim g$ dado que E .

A intuição é que eventos nulos são aqueles eventos que um agente tem certeza de que não ocorrerão. Se, e somente se, um agente tiver a certeza de que E não ocorrerá, então ele é indiferente quanto ao resultado, no caso de E , dos atos diante dele. O seguinte axioma estipula então que saber qual é o estado real não afeta a ordenação de preferência sobre **resultados**:

P3. (Neutralidade de estados)

Se $f(s_i) = X$ e $g(s_i) = Y$ sempre que $s_i \in E$ e E não é nulo, então $f \leq g$ dado que E apenas no caso de $X \leq Y$.

O próximo axioma é realmente o que torna possível determinar uma relação de crença comparativa a partir das preferências de um agente. Acima, foi sugerido que ao pedir para você apostar um prêmio sobre se uma moeda dá cara ou coroa, pode-se determinar qual desses eventos – seja cara, seja coroa – você acha mais provável. Mas essa sugestão só é plausível se o tamanho do prêmio não afetar o seu julgamento sobre a probabilidade relativa desses dois eventos. Essa suposição é capturada pelos próximos axiomas. Uma vez que o axioma é bastante complicado, ele será declarado em forma de tabela:

P4. Considere os seguintes atos:

	E	$\neg E$
f	X	X'
g	S	Y'

	F	$\neg F$
f'	X	X'
g'	S	Y'

Agora suponha que:

$$\begin{aligned} X' &\leq X, \\ Y' &\leq Y, \\ f' &\leq f \end{aligned}$$

Logo:
 $g' \leq g$.

Menos formalmente (e enunciado em termos de preferência estrita): se você prefere apostar o prêmio X em f em vez de f' , você deve considerar E mais provável que F . Portanto, você deve preferir apostar o prêmio Y em g em vez de g' uma vez que o prêmio em si não afeta a probabilidade dos eventos.

O próximo axioma não é indiscutivelmente um requisito de racionalidade, mas um dos "axiomas estruturais" de Savage (*vide* SUPPES. 2002). Um agente precisa ter alguma variação na preferência para que seja possível ler suas crenças comparativas a partir de suas preferências; e, mais geralmente, para que seja possível representá-la como maximizando a utilidade esperada. Para isso, o próximo axioma requer simplesmente que haja algumas alternativas entre as quais o agente não fique indiferente:

P5.

Há algum $f, g \in F$, tal que $f < g$.

Quando estes cinco axiomas são satisfeitos, as preferências do agente dão origem a uma relação de crença comparativa, $\preceq$, que tem a propriedade de ser uma relação de **probabilidade qualitativa** que é necessária para se poder representar $\preceq$ por uma função de probabilidade. Em outras palavras, $\preceq$ satisfaz as seguintes três condições, para quaisquer eventos E, F e G :

1. $\preceq$ é transitiva e completa,
2. se $E \cap G = \emptyset = F \cap G$, então $E \preceq F \Leftrightarrow E \cup G \preceq F \cup G$,
3. $\emptyset \preceq E, \emptyset < S$

Ser uma relação de probabilidade qualitativa não é, entretanto, suficiente para garantir a possibilidade de representação probabilística (dados os outros

axiomas de Savage). Para garantir esta possibilidade, Savage adicionou o seguinte axioma estrutural:

P6. (Não-atomicidade)

Suponha $f < g$. Então, para qualquer $X \in O$, há uma partição finita $\{E_1, E_2, \dots, E_m\}$ de S tal que:

- $f(s_j) = X$ para qualquer $s_j \in E_j$, mas $f(s_i) = f(s_j)$ para qualquer $s_j \notin E_j$,
- $g'(s_j) = X$ para qualquer $s_j \in E_j$, mas $g'(s_i) = g(s_j)$ para qualquer $s_j \notin E_j$,
- $f' < g$ e $f < g'$.

Assim como o axioma da continuidade de vNM, a não-atomicidade implica que não importa quão ruim seja um resultado X , se g já é preferível a f , então se adicionarmos X como um dos resultados possíveis de f – construindo assim uma nova alternativa f' – g ainda será preferível à alternativa modificada, desde que a probabilidade de X seja suficientemente pequena. De fato, a não-atomicidade implica que S contém eventos de probabilidade arbitrariamente pequena. Não é muito difícil imaginar como isso poderia ser satisfeito. Por exemplo, qualquer evento F pode ser particionado em dois sub-eventos equiprováveis, de acordo com o resultado de um lance de cara ou coroa. Cada sub-evento pode ser dividido de maneira semelhante, de acordo com o resultado do segundo sorteio da mesma moeda, e assim por diante, *ad infinitum*.

Savage mostrou que sempre que estes seis axiomas são satisfeitos, a relação de crença comparativa pode ser representada por uma função de probabilidade **única**. Tendo feito isso, ele poderia contar com o teorema da representação vNM para mostrar que um agente que satisfaz todos os seis axiomas¹³ pode ser representado como maximizando a utilidade esperada relativamente a uma função de probabilidade única que representa plausivelmente as crenças dos agentes sobre os estados de coisas e uma função de utilidade cardinal que representa plausivelmente os desejos do agente para os resultados (lembre-se da afirmação

¹³ Quando o conjunto de resultados é infinito, o agente precisa satisfazer um axioma adicional. Mas, para simplificar, assumiremos resultados finitos.

do teorema de Savage acima)¹⁴. A prova do próprio Savage é bastante complicada, mas Kreps (1988) fornece uma ilustração útil dela.

Não há dúvida de que o *teorema da representação da utilidade esperada*, de Savage, é muito poderoso. Há, no entanto, duas perguntas importantes a serem feitas sobre se Savage atinge seus objetivos: 1) Savage caracteriza as *preferências racionais*, pelo menos no sentido genérico? E (2) o teorema de Savage nos diz como tomar decisões racionais no mundo real? A teoria de Savage tem problemas para atender a essas duas demandas consideradas em conjunto. Pode-se arguir que a fraqueza central da teoria é que suas várias restrições e suposições puxam em diferentes direções quando se trata de construir modelos de decisão realistas e, além disso, pelo menos uma restrição (notadamente, o Princípio da Coisa Segura) só é plausível sob premissas de modelagem de decisão que supostamente são o resultado, e não o *input*, da teoria.

Um requisito bem reconhecido de modelagem de decisão para a teoria de Savage é que os resultados sejam tão específicos quanto possível em todos os aspectos importantes para sua avaliação. Se não fosse assim, o axioma da Neutralidade de Estado, por exemplo, seria uma restrição de racionalidade muito implausível. Suponha que estejamos, por exemplo, nos perguntando se vamos comprar chocolate quente ou limonada para o fim de semana, e suponha que o quão bom achamos cada opção depende de como estará o clima. Então precisamos descrever os resultados de forma que eles incluam o estado meteorológico. Pois se não o fizermos, a desejabilidade dos resultados dependerá de qual estado efetivamente ocorre. Como a limonada é, suponhamos, melhor nos dias quentes do que nos frios, um resultado como “Eu bebo limonada neste fim de semana” seria

¹⁴ Como alguns apontaram (por exemplo, ZYNDA, 2000), a função de probabilidade derivada sobre estados de coisas e a função de utilidade sobre resultados não precisam representar as crenças e desejos reais do agente. Talvez o agente tenha preferências sobre atos que satisfaçam os axiomas de Savage, que são bastante desconectados de suas crenças e desejos reais. Além disso, o teorema de Savage não exclui que um agente com preferências que satisfazem seus axiomas possa ser representado usando uma forma funcional alternativa que melhor corresponda à sua psicologia de tomada de decisão. Essas são leituras alternativas do teorema. No entanto, a inferência para a melhor explicação parece apontar em favor de considerar a representação de utilidade esperada como descobrindo crenças e desejos determinantes das preferências do agente.

mais ou menos desejável dependendo de sua ocorrência em um dia quente ou frio. Isto seria contrário ao axioma da Neutralidade de Estado. Portanto, os resultados apropriados neste caso são da forma “Eu bebo limonada neste fim de semana com tempo quente”. É claro que este resultado deve ser dividido em resultados ainda mais finos, se ainda houver outras características que afetem a escolha em questão, como compartilhar a bebida com um amigo que ama limonada *versus* compartilhar a bebida com um amigo que ama chocolate quente, e assim por diante.

O fato de que os resultados no caso acima devem ser suficientemente específicos para conter o estado da meteorologia pode parecer bastante inócuo. Entretanto, esta exigência exacerba o problema acima mencionado que muitas das opções/ações que Savage requer para seu teorema de representação são absurdas, na medida em que o conteúdo semântico dos pares estado/resposta é contraditório. Lembre-se que o domínio da ordenação de preferências na teoria de Savage equivale a **toda** função desde o conjunto de estados até o conjunto de resultados (o que Broome (1991a) chama de *Assunção de Campo Retangular*). Então, se “eu bebo limonada neste fim de semana em clima quente” é um dos resultados com os quais estamos trabalhando, e dividimos o conjunto de estados de acordo com o tempo, então deve haver, por exemplo, um ato que tenha esse resultado num estado em que faz frio! Quanto mais detalhados os resultados (como requerido para a plausibilidade da Neutralidade do Estado), menos plausível é a Assunção de Campo Retangular. Esta é uma tensão interna no quadro de Savage. Na verdade, é difícil ver como ou por que um agente racional pode/deve formar preferências sobre atos sem sentido (contudo, veja Dreier 1996 para um argumento de que esta não é uma questão tão importante). Sem esse pressuposto, porém, o ordenamento de preferências do agente não será suficientemente rico para as restrições de racionalidade de Savage produzirem o resultado da representação da UE¹⁵.

O axioma na teoria de Savage que tem recebido mais atenção é o Princípio da Coisa Segura. Não é difícil ver que este princípio entra em conflito com as preferências de Allais pela mesma razão por que estas preferências entram em conflito com a Independência (lembre-se da seção 2.3). O desafio de Allais será discutido novamente mais tarde. Por enquanto, nossa preocupação é antes o

¹⁵ No entanto, um resultado semelhante ao de Savage pode ser obtido sem a Suposição de Campo Retangular, em particular adicionando alguma estrutura extra ao conjunto de prospectos. (vide BRADLEY, 2007).

Princípio da Coisa Segura em relação à lógica interna da teoria de Savage. Para começar, o Princípio da Coisa Segura, assim como a Neutralidade do Estado, exacerba as preocupações com a Assunção de Campo Retangular. Isto porque o Princípio da Coisa Segura só é plausível se os resultados forem suficientemente específicos para responder por qualquer tipo de dependência entre os resultados em diferentes estados do mundo. Por exemplo, se o fato de se poder ter escolhido uma alternativa sem risco – e assim garantir um resultado aceitável – faz uma diferença para a conveniência de não receber nada depois de ter assumido um risco (como no problema de Allais), então isso tem que ser contabilizado na descrição dos resultados. Mas, novamente, se levarmos em conta tais dependências na descrição dos resultados, encontramos o problema de que haverá atos na ordenação de preferência que não fazem sentido (*vide* BROOME, 1991a, cap. 5).

Há um outro problema interno com a teoria de Savage associada ao Princípio da Coisa Segura: o princípio só é razoável quando o modelo de decisão é construído de tal forma que haja independência probabilística entre os atos que um agente está considerando e os estados do mundo que determinam os resultados desses atos. Lembramos que o princípio diz que se tivermos quatro opções com a seguinte forma:

	E	$\neg E$
f	X	Z
g	Y	Z
f'	X	W
g'	Y	W

Neste caso, se g é fracamente preferível a f , g' deve ser fracamente preferível a f' . Suponha, entretanto, que exista uma dependência probabilística entre os estados do mundo e as alternativas que estamos considerando, e que nós achamos que Z é melhor que ambos X e Y , e, também, achamos W melhor que ambos X e Y . Além disso, suponha que g faz $\neg E$ mais provável do que f o faz, e f' faz $\neg E$ mais provável do que g' o faz. Então, parece perfeitamente razoável preferir g a f , mas f' a g' .

Por que a exigência de independência probabilística é problemática? Para começar, em muitas circunstâncias de decisão do mundo real, é difícil enquadrar o modelo de decisão de tal forma que os estados sejam intuitivamente probabilisticamente

independentes dos atos. Por exemplo, suponha que uma pessoa gosta de fumar e está tentando decidir se quer parar ou não. Quanto tempo ela viverá está entre as contingências que afetam a desejabilidade de fumar. Seria natural dividir o conjunto de estados de acordo com o tempo de vida da pessoa. Mas então é óbvio que as opções que ela está considerando poderiam, e sem dúvida deveriam, afetar a probabilidade de ela encontrar cada estado do mundo, pois é bem reconhecido que a expectativa de vida é reduzida pelo cigarro. Savage, portanto, exigiria uma representação alternativa do problema de decisão – os estados não se referem diretamente à expectativa de vida, mas sim à propensão fisiológica do agente para reagir de uma certa forma ao fumo.

Talvez haja sempre uma maneira de se chegar a modelos de decisão tais que os atos sejam intuitiva e probabilisticamente independentes dos estados. Mas aí reside o problema mais sério. Lembremos que Savage estava tentando formular uma maneira de determinar as crenças de um agente racional a partir de suas preferências sobre os atos, de tal modo que as crenças possam ser representadas por uma função de probabilidade. Se estamos interessados em decisões do mundo real, então os atos em questão devem ser opções reconhecíveis para o agente (o que já vimos ser questionável). Além disso, agora vemos que uma das restrições de racionalidade de Savage sobre preferências – o Princípio da Coisa Segura – só é plausível se os atos modelados forem probabilisticamente independentes dos estados. Em outras palavras, essa independência deve ser embutida no modelo de decisão se for para facilitar medidas apropriadas de crença e desejo. Mas isto é para assumir que já temos informações importantes sobre as crenças do agente cujas atitudes estamos tentando representar; isto é, que estado-partições ela considera probabilisticamente independentes de seus atos.

Os problemas acima sugerem que existe a necessidade de uma teoria alternativa de escolha sob incerteza. A teoria de Richard Jeffrey, que será discutida a seguir, evita todos os problemas que têm sido discutidos até agora. Mas, como veremos, a teoria de Jeffrey tem problemas bem conhecidos, embora não sejam problemas intransponíveis.

3.2 A teoria de Jeffrey

A teoria da utilidade esperada, de Richard Jeffrey, difere da de Savage tanto em termos dos **prospectos** (ou seja, das opções) em consideração quanto

das *restrições de racionalidade nas preferências* sobre os prospectos. A vantagem distinta da teoria de Jeffrey é que problemas de decisão no mundo real podem ser modelados exatamente como o agente os percebe; a plausibilidade das restrições de racionalidade nas preferências não depende de problemas de decisão serem modelados de uma determinada forma (e pode-se acrescentar questionamentos). Primeiramente, descrevemos os prospectos ou a configuração da decisão e a regra de utilidade esperada resultante, antes de nos voltarmos às restrições de racionalidade pertinentes às preferências e ao teorema correspondente.

Ao contrário de Savage, Jeffrey não faz distinção entre os objetos de desejo instrumental e não instrumental (atos e resultados, respectivamente) e os objetos de crença (estados do mundo). Ao contrário, Jeffrey assume que **as proposições** que descrevem os estados de coisas são objetos tanto de desejo quanto de crença. À primeira vista, isto não parece objetável: assim como podemos ter visões sobre se de fato vai chover, também podemos ter visões sobre como isso seria desejável. A parte desconfortável desse arranjo é que atos também são apenas proposições – são estados de coisas comuns sobre os quais um agente tem crenças e desejos. Assim como a pessoa tem uma ordenação de preferência sobre, digamos, possíveis cenários meteorológicos para o fim de semana, ela tem uma ordenação de preferência sobre os possíveis atos que ela pode realizar, e em nenhum dos casos o estado de coisas mais preferido é necessariamente o mais provável de ser verdade. Em outras palavras, a única coisa que define os atos como especiais é o seu conteúdo substantivo – estas são as proposições de que o agente tem o poder de escolher/tornar verdade na situação em questão. É como se o agente avaliasse suas próprias opções de atuação a partir de uma perspectiva da terceira pessoa. Se se sustenta que um modelo de decisão deve representar de forma convincente a perspectiva subjetiva do agente em questão, então este é um ponto fraco da teoria de Jeffrey – embora possa ser um ponto sem consequências relevantes¹⁶.

¹⁶ Levi (1991), por exemplo, critica a teoria de Jeffrey por envolver probabilidades sobre atos que ele afirma que não podem ser sensatamente interpretados. Joyce (2002) e Rabinowicz (2002) argumentam, no entanto, que tais probabilidades são significativas e muitas vezes necessárias para a deliberação. Levi (1991), por exemplo, critica a teoria de Jeffrey por envolver probabilidades sobre atos, que ele afirma que não podem ser sensatamente interpretados. Joyce (2002) e Rabinowicz (2002) argumentam, no entanto, que tais probabilidades são significativas e muitas vezes necessárias para a deliberação.

Antes de prosseguir, uma palavra sobre proposições pode ser útil: elas são objetos abstratos que podem ser verdadeiros ou falsos, e são comumente identificados com conjuntos de mundos possíveis. Um mundo possível pode ser pensado como uma representação abstrata de como as coisas são ou poderiam ser (STALNAKER, 1987). Consulte, também, o verbete **Mundos Possíveis**, da SEP. A proposição de que “chove no tempo t ”, por exemplo, é apenas o conjunto de todos os mundos onde chove no tempo t . E esta proposição é verdadeira se, e somente se, o mundo real é um membro do conjunto de todos os mundos onde chove no tempo t . No contexto da teoria da decisão, os mundos pertinentes são os **epistemicamente** possíveis, dado o ponto de vista do agente.

O resultado básico da teoria de Jeffrey é que a desejabilidade de uma proposição, incluindo uma representando atos, depende tanto da desejabilidade das diferentes maneiras em que a proposição pode ser verdadeira, quanto da probabilidade relativa de que ela seja verdadeira dessas respectivas maneiras. Para dizer isto mais precisamente, p , q , etc., denotarão variáveis proposicionais. Seja $\{p_1, p_2, \dots, p_n\}$ uma entre muitas partições finitas da proposição p ; ou seja, conjuntos de maneiras nas quais a proposição p pode ser realizada que são mutuamente incompatíveis, mas conjuntamente exaustivas. Por exemplo, se p é a proposição “está chovendo”, então poderíamos particionar essa proposição muito grosseiramente de acordo com se vamos à praia ou não, mas também poderíamos particionar p muito mais precisamente – por exemplo, de acordo com a quantidade precisa de chuvas em milímetros por hora. A *desejabilidade* de p , segundo Jeffrey, denotada $Des(p)$, é dada por:

Equação de Jeffrey:

$$Des(p) = \sum_i Des(p_i) \cdot P(p_i | p)$$

Esta é de fato uma fórmula de utilidade esperada *condicional* para avaliar p . Como observado, um caso especial é quando o conteúdo de p é tal que é reconhecidamente algo que o agente pode escolher tornar verdadeiro, ou seja, um ato.

Uma diferença importante entre a fórmula de desejabilidade de Jeffrey e a fórmula de utilidade esperada de Savage é que na de Jeffrey não há distinção feita entre desejabilidade e desejabilidade “esperada”, ao contrário do que deve ser feito na teoria de Savage, onde há uma distinção clara entre utilidade, que mede os desejos fundamentais de um agente em relação aos últimos resultados, e a utilidade esperada, que mede as preferências de um agente sobre prospectos ou

atos incertos. Essa desanalogia se deve ao fato de que não há nenhum sentido no qual as proposições do tipo “ p_i ”, nos termos dos quais a variável p é avaliada, necessariamente precise referir os últimos resultados; elas próprias podem ser consideradas como prospectos incertos que são avaliados em termos de suas diferentes realizações possíveis.

Outra coisa importante a se notar sobre a maneira de Jeffrey calcular a desejabilidade é que ela não assume independência probabilística entre a alternativa que está sendo avaliada, p , e as formas possíveis, as “ p_i ”, em que a alternativa pode ser realizada. De fato, a probabilidade de cada p_i é explicitamente condicional em relação ao p em questão. Quando se trata de avaliar atos, isto quer dizer (na terminologia de Savage) que as probabilidades para os possíveis pares de estados para o ato são condicionadas ao ato em questão. Assim, vemos por que a pessoa pode descrever seu problema de decisão exatamente como ela o vê; não há nenhuma exigência de que ela identifique um conjunto de estados (no caso de Jeffrey, esta seria uma partição do espaço proposicional que é ortogonal à partição do ato) de forma que os estados sejam adequadamente refinados e probabilisticamente independentes dos atos.

Além disso, deve ser evidente, dada a discussão do Princípio da Coisa Segura (PCS) na Seção 3.1, que a teoria de Jeffrey não tem esse axioma. Como os estados podem ser probabilisticamente dependentes de atos, um agente pode ser representado como maximizando o valor da função de desejabilidade de Jeffrey e ao mesmo tempo violando o PCS. Além disso, ao contrário de Savage, o teorema da representação de Jeffrey não depende de nada como a Assunção de Campo Retangular. O agente não é obrigado a ter preferências sobre atos ou proposições artificialmente construídas que se revelem sem sentido, dada a interpretação de estados e resultados particulares. Na verdade, apenas aquelas proposições que o agente considera possíveis (no sentido de que lhes atribui uma probabilidade maior que zero) estão, segundo a teoria de Jeffrey, incluídas em sua ordenação de preferências.

Claro que ainda precisamos de certos pressupostos estruturais para provar um teorema de representação para a teoria de Jeffrey. Em particular, o conjunto Ω , sobre o qual se define a ordenação de preferências $\leq$, tem de ser uma *álgebra booleana sem átomos* de proposições, da qual foram retiradas as proposições impossíveis, denotada por $\perp$. Uma álgebra booleana é apenas um conjunto de, por exemplo, proposições ou sentenças que é fechado sob os clássicos operadores lógicos (incluindo a negação). Uma álgebra é **sem átomo**, caso todos os seus elementos possam ser divididos em elementos mais granulares. A suposição de

que Ω é sem átomo é, portanto, semelhante à P6 de Savage, e lhe pode ser dada uma justificação semelhante: qualquer forma p_i na qual p pode ser verdadeiro pode ser dividida em duas outras proposições de acordo com o resultado de um “cara ou coroa”.

Então, em que condições uma relação de preferência $\leq$ no conjunto Ω pode ser representada como maximizadora da desejabilidade? Algumas das condições exigidas de preferência devem ser familiares agora e não serão discutidas mais. Em particular, $\leq$ tem que ser transitiva, completa e contínua (lembre-se de nossa discussão na Seção 2.3 do axioma de preferência de continuidade de vNM).

As próximas duas condições, no entanto, não são explicitamente parte dos dois teoremas de representação que foram considerados até agora:

Média

Se $p, q \in \Omega$ são mutuamente incompatíveis, então $p \leq q \Leftrightarrow p \leq (p \cup q) \leq q$ ¹⁷

Imparcialidade

Suponha que $p, q \in \Omega$ sejam mutuamente incompatíveis e $p \sim q$. Então, se $(p \cup r) \sim (q \cup r)$ para algum r que é mutuamente incompatível com ambos p e q , e tal que $\neg (r \sim p)$, então $p \cup r \sim q \cup r$ é verdadeiro para cada r .

A condição **Média** é a condição de racionalidade distintiva na teoria de Jeffrey. Na verdade, ela pode ser vista como uma versão fraca da Independência e do Princípio da Coisa Segura, tendo um papel semelhante na teoria de Jeffrey. Mas não é diretamente incompatível com as preferências de Allais, e sua plausibilidade não depende do tipo de independência probabilística que o PCS implica. O postulado exige que nenhuma proposta seja estritamente melhor ou pior do que todas as suas possíveis realizações, o que parece ser um requisito razoável. Quando p e q são mutuamente incompatíveis, $p \cup q$ implica que p ou q é verdade, mas não ambas. Portanto, parece razoável que $p \cup q$ deve não ser nem mais nem menos desejável que p e q . Suponha que ou p ou q é mais desejável que o outro. Então, como $p \cup$

¹⁷N.T.: “U” aqui refere a operação de “união de conjuntos”, equivalendo, pois, ao conectivo lógico “ou” - como o autor deixa claro a seguir; Jeffrey (1963) usa o sinal “V” no lugar.

q é compatível com a verdade do mais ou do menos desejável dos dois, a desejabilidade de $p \cup q$ deve estar **estritamente** entre a de p e a de q . Entretanto, se p e q são igualmente desejáveis, então $p \cup q$ deve ser tão desejável quanto cada uma das duas.

O apelo intuitivo da Imparcialidade, que desempenha na teoria de Jeffrey um papel semelhante ao de P4 na de Savage, não é tão grande quanto o da **Média**. O próprio Jeffrey admitiu o mesmo em seu comentário: “O axioma está lá porque precisamos dele, e se justifica por nossa crença anterior na plausibilidade do resultado que pretendemos deduzir dele” (JEFFREY, 1965, p. 147).

No entanto, parece que se pode argumentar que qualquer pessoa razoável irá satisfazer este axioma. Suponha que você fique indiferente entre duas proposições, p e q , que não podem ser simultaneamente verdadeiras. E suponha agora que encontramos uma proposição r , que é incompatível com ambos p e q , e que você acha mais desejável do que ambos p e q . Então, se se verificar que você é indiferente entre “ p em conjunto com r ” e “ q em conjunto com r ”, isso deve ser porque você acha p e q igualmente prováveis. Caso contrário, você preferiria a conjunção contendo aquele entre p e q que você acha menos provável, já que isso lhe dá uma chance maior da proposta mais desejável r . Segue-se que, para qualquer outra proposição s que satisfaça as condições r acima mencionadas, você também deve ficar indiferente entre $p \cup s$ e $q \cup s$, já que, novamente, as duas conjunções são igualmente prováveis de resultar em s .

A primeira pessoa a provar um teorema afirmando condições suficientes para que uma relação de preferência fosse representativa como maximização do valor de uma função de desejabilidade de Jeffrey não foi na verdade o próprio Jeffrey, mas o matemático Ethan Bolker (1966, 1967). Ele provou o seguinte resultado (lembre-se da definição de “medida de desejabilidade” dada acima)¹⁸:

Teorema 4 (Bolker)

Seja Ω uma álgebra de proposições booleana completa e sem átomos, e $\leq$ uma relação contínua, transitiva e completa em $\Omega \setminus \perp$, que satisfaça a **Média** e a **Imparcialidade**. Então há uma medida de

¹⁸ A prova de Bolker é bastante complicada e parece não existir nenhuma introdução para estudantes.

desejabilidade em $\Omega \setminus \perp$ e uma medida de probabilidade em Ω em relação à qual $\leq$ pode ser representada como maximizando a desejabilidade.

Infelizmente, o teorema da representação de Bolker não produz um resultado nem de perto tão único quanto o de Savage. Isto é, mesmo que as preferências de uma pessoa satisfaçam todas as condições do teorema de Bolker, não é garantido que haverá apenas uma função de probabilidade que represente suas crenças, nem que a função de desejabilidade que representa seus desejos será única até uma transformação linear positiva. Pior ainda, a mesma ordenação de preferências que satisfaz todos esses axiomas poderia ser representada como maximizando a desejabilidade em relação a duas funções de probabilidade que nem sequer concordam em como ordenar as propostas de acordo com sua probabilidade¹⁹.

Para aqueles que pensam que a única maneira de determinar as crenças comparativas de uma pessoa é olhar para suas preferências, a falta de singularidade na teoria de Jeffrey é um grande problema. Na verdade, esta pode ser uma das principais razões pelas quais os economistas ignoraram em grande parte a teoria de Jeffrey. Os economistas têm sido tradicionalmente céticos em relação a qualquer conversa sobre os desejos e crenças de uma pessoa que vá além do que pode ser estabelecido pelo exame das preferências da pessoa, que eles consideram ser a única atitude que é revelada diretamente pelo comportamento de uma pessoa. Para esses economistas, portanto, não é uma notícia bem-vinda não podermos, em princípio, determinar as crenças comparativas de uma pessoa racional através do exame das suas preferências.

Aqueles que estão menos inclinados ao behaviorismo podem, no entanto, não achar essa falta de singularidade no teorema de Bolker como um problema. James Joyce (1999), por exemplo, acha que a teoria de Jeffrey acerta exatamente nesse sentido, já que não se deve esperar que condições razoáveis impostas às preferências de uma pessoa sejam suficientes para determinar uma função de

¹⁹ Isso realmente decorre do fato de que as preferências não determinam uma única função de probabilidade, dado que o domínio da função de probabilidade é infinito (esse último ponto decorre da ausência de atomização de Ω). Pois se P e Q são funções de probabilidade em um domínio infinito que concordam com a ordem de todas as proposições em seu domínio, P e Q são a mesma função de probabilidade (VILLEGAS, 1964).

probabilidade única que represente as crenças da pessoa. Só impondo condições excessivamente fortes, como faz Savage, é que podemos conseguir isso. Entretanto, se a singularidade é o que procuramos, então podemos, como Joyce aponta, complementar os axiomas Bolker-Jeffrey com certas condições na relação de crença comparativa do agente que garantem que ela dê origem a uma e apenas uma função de probabilidade, como, por exemplo, as condições de crença comparativa propostas em Villegas (1964). Feito isso, podemos mostrar que qualquer função que represente os desejos do agente será única até uma transformação linear positiva.

Ao invés de acrescentar postulados específicos de crenças à teoria de Jeffrey, como Joyce sugere, pode-se obter o mesmo resultado de singularidade enriquecendo o conjunto de prospectos. Richard Bradley (1998), por exemplo, mostrou que se a álgebra booleana na teoria de Jeffrey for estendida aos condicionais indicativos, então uma relação de preferência no domínio estendido que satisfaça os axiomas Bolker-Jeffrey (e alguns axiomas relacionados que se aplicam especificamente aos condicionais) será representável como maximizando a desejabilidade, onde a função de probabilidade e a função de desejabilidade são únicas até uma transformação linear positiva.

4. Implicações mais amplas da teoria da Utilidade Esperada (UE)

Observou-se desde o início que a teoria da UE é tanto uma teoria da escolha racional, ou de preferências gerais entre atos, quanto uma teoria da crença e do desejo racionais. Esta seção expande, por sua vez, os compromissos epistemológicos e avaliativos da teoria da UE.

4.1 Sobre a crença racional

Alguns referem-se à teoria da UE como *teoria da decisão bayesiana*. Este rótulo traz à tona o compromisso com o **probabilismo**, ou seja, que as crenças podem vir em graus que, sob pena de irracionalidade, podem ser representados numericamente de acordo com o cálculo de probabilidade. Portanto, há uma forte conexão entre a teoria da UE e o probabilismo, ou mais geralmente, entre a preferência racional e a crença racional (os detalhes mais sutis sobre preferência racional, e a crença racional correspondente, não são o foco aqui; os desafios à

teoria da UE sobre este aspecto são abordados na Seção 5 abaixo).

Alguns consideram que a conexão entre preferência racional e crença racional é realmente muito profunda. No extremo oposto do espectro está a posição de que o próprio significado da crença envolve preferência. Aliás, lembre-se dessa manobra na teoria de Savage discutida anteriormente na Seção 3.1. No entanto, muitos questionam a plausibilidade de equiparar a crença comparativa com preferências sobre prospectos especialmente planejados. Uma posição mais moderada é considerar essas preferências como implicadas por crenças comparativas relevantes, mas não idênticas a elas. Sejam as crenças meramente fundamentadas, sejam definidas em termos de preferência, há uma outra questão sobre se a única justificativa para a crença racional que tem uma certa estrutura (digamos, em conformidade com o cálculo de probabilidade) é pragmática; ou seja, um argumento que se baseia nas preferências do agente seria, de outro modo, inconsistente ou autodestrutivo (*self-defeating*). Um defensor recente desse tipo de pragmatismo (embora expresso em termos mais gerais) é Rinard (2017). Outros afirmam que as concepções de crença racional podem e devem ser justificadas em termos epistêmicos; Joyce (1998), por exemplo, oferece uma justificativa não pragmática do probabilismo que se baseia na noção de “distância em relação à verdade” das crenças de alguém (para mais desenvolvimentos desta posição, veja o verbete sobre **argumentos de utilidade epistêmica em favor do probabilismo**).

Apesar dessas disputas mais sutis, os bayesianos concordam que as considerações pragmáticas desempenham um papel significativo no gerenciamento das crenças. Pelo menos uma maneira importante pela qual um agente pode investigar seus graus de crença é refletir sobre suas implicações pragmáticas. Além disso, buscar ou não mais evidências é uma questão pragmática; depende do “valor da informação” que se espera obter com relação ao problema de decisão em questão. A ideia é que buscar mais evidências é uma ação que vale a pena escolher apenas no caso de a utilidade esperada de buscar mais evidências (antes de tomar uma decisão) ser maior do que a utilidade esperada de tomar a decisão com base nas evidências existentes. Esse raciocínio ganhou destaque em um artigo de Good (1967), onde ele prova que sempre se devem buscar “evidências gratuitas” que possam ter influência na decisão em questão (os precursores deste teorema podem ser encontrados em Ramsey 1990, publicado postumamente, e Savage 1954). Observe que o teorema assume a regra de aprendizado Bayesiana padrão conhecida como “condicionalização”, que exige que quando a experiência de aprendizagem de alguém tenha a forma de vir a conhecer com certeza alguma proposição (à qual

se tem atribuído probabilidade positiva), os novos graus de crença desse alguém devem ser iguais aos antigos graus de crença condicionais à proposição que agora tem probabilidade 1. Com efeito, o fato de que a condicionalização desempenha um papel crucial no resultado de Good sobre o valor não negativo da evidência gratuita é considerado por alguns como uma justificativa para essa regra de aprendizagem.

Portanto, a teoria da UE, ou teoria da decisão bayesiana, sustenta um poderoso conjunto de normas epistêmicas. Ela tem sido considerada a concepção apropriada de inferência científica, dando origem a uma escola de inferência estatística e *design* de experiências, e convidando interpretações formais de conceitos-chave como "evidência", "suporte probatório", "indução" versus "abdução" e a influência de "coerência" e "poder explicativo" sobre a verdade (veja os verbetes relevantes de tópicos correlatos na SEP). O principal concorrente do bayesianismo, no que diz respeito à inferência científica, é indiscutivelmente o conjunto de abordagens conhecidas como Estatística Clássica ou Estatística de Erros, que negam o sentido de "graus de suporte" (probabilísticos ou não) conferidos por evidências a uma hipótese. Essas abordagens enfocam, em vez disso, na questão de se uma hipótese sobreviveu a vários "testes severos", e as inferências são feitas tendo em vista as propriedades de longo prazo dos testes em oposição a como eles funcionam em qualquer caso único, o que exigiria raciocínio teórico da decisão (veja o verbete sobre **filosofia da estatística**, na SEP).

4.2 Sobre o desejo racional

A teoria da UE também toma uma posição sobre a estrutura do desejo racional. A este respeito, a teoria tem sido criticada em frentes opostas. Consideremos, primeiro, a crítica de que a teoria da UE é muito permissiva em relação ao que pode influenciar os desejos de um agente. Voltemo-nos, então, para a crítica oposta: que, quando se trata de desejo, a teoria da UE não é permissiva o suficiente.

A preocupação de que a teoria da UE seja muito permissiva em relação ao desejo está relacionada à preocupação de que a teoria seja *não falseável*. A preocupação é que preferências aparentemente irracionais, pelas luzes da teoria da UE, possam sempre ser interpretadas como racionais, sob uma descrição adequada das opções em estudo. Como discutido na Seção 1, acima, as preferências que parecem violar a Transitividade podem ser interpretadas como consistentes com este axioma, desde que as opções a serem comparadas variem em sua

descrição, dependendo, entre outras coisas, das outras opções em consideração. Isso também vale para preferências que parecem violar a Separabilidade ou Independência (da contribuição de cada resultado para o valor global de uma opção), discutidas ainda na Seção 5.1 abaixo. Pode-se argumentar que esta é a maneira correta de descrever as preferências desses agentes. Afinal, um modelo adequado de preferência é supostamente aquele que captura, na descrição de últimos resultados e de opções, tudo o que importa para um agente. Nesse caso, no entanto, a teoria da UE é efetivamente vazia ou impotente como um padrão de racionalidade a que os agentes podem aspirar. Além disso, estende a noção de o que são propriedades genuínas de resultados que podem, razoavelmente, conferir valor ou ser desejáveis para um agente.

Há duas maneiras de reagir à ideia de que as preferências de um agente são necessariamente consistentes com a teoria da UE, com as implicações acima mencionadas para o que o agente pode desejar:

- Pode-se resistir à alegação, afirmando que há restrições adicionais no **conteúdo** das preferências de um agente. Por um lado, pode haver restrições empíricas pelas quais o conteúdo das preferências é determinado por alguma ponderação entre adequação e simplicidade na representação da “rede” mais ampla de atitudes de preferência do agente. Por outro lado, pode haver restrições normativas com relação a quais tipos de resultados um agente pode razoavelmente discriminar. Para uma discussão relevante, confira os trabalhos de Tversky (1975), Broome (1991a, 1993), Pettit (1993), Dreier (1996), Guala (2006) e Vredenburg (2020).
- Pode-se, alternativamente, abraçar a reivindicação, interpretando a teoria da UE não como um padrão contra o qual um agente pode passar ou falhar, mas sim como um princípio organizador que permite a caracterização dos desejos de um agente, bem como suas crenças (*vide* GUALA, 2008).

De qualquer forma, pode-se argumentar ainda que a teoria da UE não vai longe o suficiente ao estruturar as atitudes de preferência de um agente de tal modo que possamos entender as **razões** dessas atitudes de preferência. Dietrich e List

(2013, 2016a) propuseram um quadro mais geral para preencher essa lacuna. Em seu quadro, as preferências que satisfazem algumas restrições mínimas são representáveis enquanto dependentes do pacote de propriedades em termos dos quais cada opção é percebida pelo agente em um determinado contexto. As propriedades podem, por sua vez, ser categorizadas como **propriedades de opção** (que são intrínsecas ao desfecho), propriedades **relacionais** (que dizem respeito ao desfecho em um contexto específico) ou propriedades de **contexto** (que dizem respeito ao contexto da própria escolha). Tal representação permite uma análise mais detalhada das razões das preferências de um agente e captura diferentes tipos de dependência de contexto nas escolhas de um agente. Além disso, ela permite restrições explícitas sobre o que conta como uma razão legítima de preferência, ou, em outras palavras, quais propriedades legitimamente apresentam em uma descrição do resultado; tais restrições podem ajudar a esclarecer os compromissos normativos da teoria da UE.

Há, também, modelos menos gerais que oferecem exemplos para entender as razões subjacentes às preferências. Por exemplo, a estrutura (*framework*) de decisão de múltiplos critérios (*vide* KEENEY; RAIFFA, 1993) considera a ordenação geral de preferências de um agente sobre opções como um agregado do conjunto de ordenações de preferências a qual correspondente a todas as dimensões pertinentes de valor. Sob certas suposições, a ordenação geral ou agregada de preferências é compatível com a teoria da UE. Pode-se procurar entender o papel do tempo, ou a posição temporal dos bens, sobre as preferências. Para isso, os resultados são descritos em termos de feixes (*bundles*) de bens temporalmente indexados, ou fluxos de consumo. Para um modelo inicial deste tipo, confira Ramsey (1928); um tratamento posterior mais influente é o de Koopmans (1960). Pode haver uma estrutura sistemática para as preferências de um agente em relação a esses fluxos de consumo, além da estrutura imposta pelos axiomas de preferência da UE. Por exemplo, os autores supracitados consideraram e caracterizaram preferências que exibem **descontos temporais exponenciais**.

Vamos agora ao tipo de crítica oposta: que as restrições limitadas que a teoria da UE impõe à preferência e ao desejo racionais são, por outro lado, excessivamente restritivas. Aqui, o foco será a compatibilidade da teoria da UE com posições éticas proeminentes em relação à escolha dos atos, bem como posições metaéticas quanto à natureza do valor e à sua relação com a crença.

Pode-se muito bem perguntar se a teoria da UE, e na verdade a teoria da decisão em geral, é neutra no que diz respeito à ética normativa, ou se é compatível

apenas com o **consequencialismo ético**, dado que a classificação de um ato é totalmente determinada pela utilidade de seus possíveis resultados. Tal modelo parece estar em desacordo com **teorias éticas não consequencialistas**, para as quais a escolha de atos supostamente depende de outras coisas além do valor moral de suas consequências. O modelo não parece capaz de acomodar noções deontológicas básicas como a relatividade de agentes, proibições absolutas, ou atos permissíveis, mas sub-ótimos.

Uma resposta inicial, no entanto, é que não se deve extrapolar demais a partir dos conceitos formais da teoria da decisão. A medida de utilidade sobre atos e resultados é simplesmente uma maneira conveniente de representar uma ordenação, e deixa muito espaço para diferentes maneiras de identificar e avaliar resultados. Assim como a função de utilidade de um agente não precisa ser insensível às considerações éticas em geral (um equívoco comum devido à prevalência de preferências egoístas em modelos econômicos; consulte o trabalho de Sen, de 1977), ela também precisa ser insensível a considerações éticas especificamente deontológicas ou não consequencialistas. Tudo depende de como os atos e seus resultados são distinguidos e avaliados. Para começar, o caráter de um ato pode ser uma propriedade de todos os seus resultados possíveis. Além disso, pode ser relevante se um evento ocorre, ou se é perpetrado especificamente pelo agente tomador de decisão ou outra pessoa. O fato de um ato envolver mentira, digamos, pode ser referenciado em todos os resultados possíveis do ato e, além disso, essa mentira por parte do agente de decisão pode ser distinguida da mentira de outros. Em geral, os atos e seus resultados podem ser distinguidos de acordo com o que quer que seja moral, seja isso uma propriedade relacional complexa associada a como e quando o ato é escolhido, por quem e/ou de que forma algum estado de coisas resulta do ato. Para discussões iniciais sobre como uma ampla gama de propriedades éticas pode ser acomodada na descrição de atos e resultados, consulte, por exemplo, Sen (1982), Vallentyne (1988), Broome (1991b) e Dreier (1993). Desde então, essa ideia tem sido adotada por outros associados ao chamado programa de “consequencialização”, incluindo Louise (2004) e Portmore (2007). A ideia é que o conselho normativo de teorias éticas supostamente não consequencialistas pode ser representado em termos de uma classificação de atos/resultados correspondentes a alguma função de valor, como ocorre com teorias éticas consequencialistas (ver também Colyvan et al. 2010).

Um ponto crítico para reconciliar a teoria da decisão com todas as formas de não consequencialismo é a dificuldade em acomodar **proibições absolutas** ou

condições restritivas (*side-constraints*), confira Oddie e Milne (1999), Jackson e Smith (2006). Por exemplo, suponha que haja uma proibição moral de matar uma pessoa inocente, não importa o que esteja em jogo. Talvez tal restrição seja melhor modelada em termos de uma classificação lexical e da função de valor correspondente, na qual o status de “matar inocentes” de um ato ou resultado acaba tendo prioridade na determinação de sua classificação ou valor relativo. Mas isso tem implicações contraintuitivas em face do risco, uma vez que muitos atos terão alguma chance, embora pequena, de matar um inocente. A lição aqui pode ser simplesmente que as teorias em questão requerem desenvolvimento; qualquer teoria ética madura nos deve uma explicação de como agir sob risco ou incerteza. O que é indiscutivelmente um desafio mais convincente para reconciliar a teoria da decisão com o não consequencialismo é a acomodação de “opções centradas no agente” e a “supererrogação” associada. Portmore (2007) e Lazar (2017) oferecem propostas nesse sentido, as quais apelam (de maneiras diferentes) para a classificação moral de atos/resultados como distintos dos custos pessoais para o agente de perseguir esses atos/resultados.

Na medida em que a teoria da decisão pode ser reconciliada com toda a gama de teorias éticas, deveríamos dizer que não há distinções significativas entre essas teorias? Brown (2011) e Dietrich e List (2017) demonstram que, de fato, a representação teórica da escolha de teorias éticas facilita melhor as distinções entre elas; termos como “(não) consequencialismo” podem ser definidos com precisão, embora de maneiras discutíveis. De forma mais geral, podemos catalogar teorias em termos dos tipos de propriedades (sejam intrínsecas ou em algum sentido relacionais) que distinguem atos/resultados e, também, em termos da natureza da classificação de atos/resultados que eles produzem (seja transitiva, completa, contínua e assim por diante). Isso também serve para revelar desvios da teoria da UE.

Na verdade, alguns dos contraexemplos mais convincentes aos axiomas de preferência da UE baseiam-se em considerações éticas. Lembre-se de nossa discussão anterior sobre os axiomas básicos de ordenação na Seção 1. O axioma da transitividade foi desafiado pelo apelo a exemplos eticamente motivados de ciclos de preferência (*vide* TEMKIN, 2012). A noção de uma ordenação lexical não contínua foi mencionada acima em relação às condições restritivas éticas. A dispensabilidade do axioma da Completude também é frequentemente motivada pelo apelo a exemplos envolvendo valores éticos concorrentes entre os quais é difícil ponderar, como bem-estar médio *versus* total. Outros exemplos sugestivos contra a Completude envolvem noções concorrentes de bem-estar pessoal (*vide*

LEVI, 1986; CHANG, 2002). Um agente racional deve ter uma preferência definida entre, digamos, duas opções de carreira que pesam em direções diferentes no que diz respeito às oportunidades de autoexpressão criativas *versus* serviço comunitário (talvez uma carreira na dança *versus* uma carreira praticando medicina em regiões remotas)? Observe que alguns desses desafios à teoria da UE são discutidos com mais profundidade na Seção 5 abaixo.

Finalmente, voltamo-nos para os potenciais compromissos meta-éticos da teoria da UE. David Lewis (1988, 1996) empregou a teoria da UE para argumentar contra o **anti-Humeanismo**, a posição de que às vezes somos movidos inteiramente por nossas crenças sobre o que seria bom, ao invés de sermos por nossos desejos, como afirmam os **Humeanos**. Ele formulou a teoria anti-Humeana como postulando uma conexão necessária entre, por um lado, o desejo de um agente em relação qualquer proposição A, e, por outro lado, sua crença em uma proposição sobre o quanto A é bom; e alegou provar que quando tal conexão é formulada em termos da teoria da UE, o agente em questão será dinamicamente incoerente. Várias pessoas criticaram o argumento de Lewis. Por exemplo, Broome (1991c), Byrne e Hájek (1997) e Hájek e Pettit (2004) sugerem formulações do anti-Humeanismo que são imunes às críticas de Lewis, enquanto Stefánsson (2014) e Bradley e Stefánsson (2016) argumentam que a demonstração de Lewis se baseia em uma suposição falsa. Não obstante, o argumento de Lewis sem dúvida provocou um debate interessante sobre os tipos de conexões entre crença e desejo que a teoria da UE permite. Além disso, existem outras questões de relevância metaética que se pode investigar a respeito do papel e da estrutura do desejo na teoria da UE. Por exemplo, Jeffrey (1974) e Sen (1977) oferecem algumas investigações preliminares sobre se a teoria pode acomodar desejos/preferências de **ordem superior** e, em caso afirmativo, como eles se relacionam com os desejos/preferências de primeira ordem.

5. Desafios à teoria da UE

Até o momento o foco tem sido em versões proeminentes da teoria padrão de escolha racional: a teoria da UE. Esta seção retoma as principais críticas à teoria da UE que foram desenvolvidas em concepções alternativas de escolha racional. As inovações propostas para a teoria padrão são distintas e, por isso, são discutidas separadamente, mas não são necessariamente mutuamente exclusivas. Observe que não abordamos todas as críticas à teoria da UE que inspiraram relatos alternativos

de escolha racional. Duas omissões principais deste tipo (por falta de espaço e, também, porque foram exaustivamente abordadas em outros verbetes da SEP) são: i) o problema das anomalias causais e o desenvolvimento da teoria da decisão causal (ver o verbete sobre a teoria da decisão causal na SEP) e ii) o problema dos espaços de estado infinito e o desenvolvimento de alternativas como a “teoria da expectativa relativa” (veja os verbetes sobre **teorias normativas da escolha racional: teoria da utilidade esperada e o paradoxo de São Petersburgo**).

5.1 Sobre risco e atitudes de arrependimento

A teoria da utilidade esperada tem sido criticada por não ensejar interações de valor entre resultados em estados do mundo distintos e mutuamente incompatíveis. Por exemplo, lembre-se que ao decidir entre duas opções de risco você deve, de acordo com a versão da teoria de Savage, ignorar os estados do mundo onde as duas opções resultam no mesmo desfecho. Isso parece muito razoável se pudermos assumir a *separabilidade* entre resultados em diferentes estados do mundo; isto é, se a contribuição que um resultado em um estado do mundo faz para o valor global de uma opção for independente de quais outros desfechos a opção possa resultar. Pois, então, resultados idênticos (com probabilidades iguais) devem cancelar um ao outro numa comparação de duas opções, o que implicaria que se duas opções compartilham um resultado em algum estado do mundo, então ao comparar as opções, não importa o que esse resultado compartilhado é.

O paradoxo de Allais, discutido na Seção 2.3 acima, é um exemplo clássico onde a separabilidade anteriormente mencionada falha. Para facilitar a compreensão, as opções que geram o paradoxo são reproduzidas na Tabela 3. Lembre-se da Seção 2.3 na qual vimos que as pessoas tendem a preferir L2 a L1, e L3 a L4 – uma atitude que foi chamada *preferência de Allais*, em violação à teoria da utilidade esperada. A violação ocorre precisamente porque as contribuições que alguns desses resultados fazem para o valor geral de uma opção não são independentes dos outros resultados que a opção pode ter. Compare a chance extra de resultado \$0 que L1 tem sobre L2 com a mesma chance extra de \$0 que L3 tem sobre L4. Muitas pessoas pensam que essa chance extra conta mais fortemente na primeira comparação do que na última, ou seja, que uma chance extra de probabilidade 0,01 de \$0 contribui com um valor negativo maior para L1 do que para L3. Alguns explicam isso apontando que o **arrependimento** que alguém experimentaria por não ganhar

nada, quando se poderia ter \$ 2.400 com certeza – ou seja, ao escolher L1 em vez de L2 no caso de o primeiro bilhete ser sorteado – é muito maior do que o arrependimento que alguém experimentaria ao não ganhar nada quando a opção rejeitada também tinha uma grande chance de resultar em \$0 – como ao escolher L3 em vez de L4 (*vide* LOOMES; SUGDEN, 1982). Mas seja a preferência em questão explicada ou não pelo potencial para arrependimento, parece que a desejabilidade do resultado \$0 deveria depender do que poderia ser (ou teria sido) de outra forma – violando assim o pressuposto de separabilidade acima mencionado. Confira Thoma (2020a) para uma ampla discussão recente dessa suposição.

	1	2-34	35-100
L ₁	\$ 0	\$ 2.500	\$ 2.400
L ₂	\$ 2.400	\$ 2.400	\$ 2.400

	1	2-34	35-100
L ₃	\$ 0	\$ 2.500	\$ 0
L ₄	\$ 2.400	\$ 2.400	\$ 0

Tabela 3. Paradoxo de Allais

Várias tentativas têm sido feitas para tornar as preferências de Allais compatíveis com alguma versão da teoria da utilidade esperada. Uma resposta comum é sugerir que o problema de escolha foi descrito incorretamente. Se é realmente racional avaliar \$0 de forma diferente dependendo de qual loteria faz parte, então talvez isto deva ser contabilizado na descrição dos resultados (BROOME, 1991a). Por exemplo, poderíamos adicionar uma variável ao resultado de \$0 que poderia resultar em representar o arrependimento ou risco extra associado a esse resultado em comparação com os resultados de \$0 das outras loterias (como feito na Tabela 4). Se fizermos isso, as preferências de Allais não são mais inconsistentes com a teoria da UE. A maneira mais simples de ver isso é notar que quando ignoramos o estado do mundo onde as opções que estão sendo comparadas têm o mesmo resultado (ou seja, quando ignoramos a última coluna da Tabela 4), L1 não é mais idêntico a L3, o que significa que o axioma de Independência de von Neumann e Morgenstern (e o Princípio da Coisa Segura de Savage) não exige mais que se prefira L2 a L1 apenas caso se prefira L4 a L3.

	1	2-34	35-100
L_1	\$ 0 + δ	\$ 2.500	\$ 2.400
L_2	\$ 2.400	\$ 2.400	\$ 2.400

	1	2-34	35-100
L_3	\$ 0	\$ 2.500	\$ 0
L_4	\$ 2.400	\$ 2.400	\$ 0

Tabela 4. Paradoxo de Allais re-descrito

A “estratégia de redescrição” acima pode ser empregada sempre que o valor e/ou contribuição de um resultado depender de outros resultados possíveis: apenas descreva os resultados de uma forma que explique esta dependência. Porém, o mais preocupante é que a estratégia pode ser empregada sempre que nos depararmos com **qualquer** violação da teoria da utilidade esperada ou de outras teorias da racionalidade (conforme discutido na Seção 4.2).

Lara Buchak (2013) desenvolveu recentemente uma teoria da decisão que pode acomodar as preferências de Allais sem a redescrição dos resultados. Na interpretação de Buchak, a explicação para as preferências de Allais não é o **valor** diferente que o resultado \$0 tem, dependendo de qual loteria ele compõe. O resultado em si tem o mesmo valor. No entanto, a **contribuição** que \$0 faz para o valor geral de uma opção depende parcialmente de quais outros resultados são possíveis, sugere ela, o que reflete o fato de que o risco da opção que a possibilidade de \$0 gera depende de em quais outros desfechos a opção pode resultar. Para acomodar as preferências de Allais (e outras atitudes de risco intuitivamente racionais que violam a teoria da UE), Buchak introduz uma **função de risco** que representa a disposição das pessoas em trocar as chances de algo bom pelos riscos de algo ruim. E ela mostra que se um agente satisfaz um conjunto particular de axiomas, que é essencialmente o de Savage (exceto que o Princípio da Coisa Certa é substituído por um axioma estritamente mais fraco), então as preferências do agente podem ser representadas como maximizando a **utilidade esperada ponderada pelo risco** – que é essencialmente a utilidade esperada do estilo Savage, ponderada por uma função de risco.

Bradley e Stefánsson (2017) também desenvolveram uma nova teoria da decisão, em parte, em resposta ao paradoxo de Allais. Mas, ao contrário de Buchak, eles sugerem que o que explica as preferências de Allais é que o valor de não ganhar nada em uma loteria escolhida depende parcialmente do que teria acontecido se o agente tivesse escolhido de forma diferente. Para acomodar isso, eles estendem a álgebra booleana na teoria da decisão de Jeffrey para proposições **contrafactuais** e mostram que a teoria estendida de Jeffrey pode representar as dependências de valor que muitas vezes encontramos entre resultados contrafactuais e reais. Em particular, a teoria deles pode capturar a intuição de que a (in)desejabilidade de não ganhar nada depende, em parte, de se alguém tinha ou não garantia de ganhar algo caso houvesse escolhido de forma diferente. Portanto, sua teoria pode representar as preferências de Allais como maximizando o valor de uma função de desejabilidade de Jeffrey estendida.

Stefánsson e Bradley (2019) sugerem ainda outra maneira de explicar as preferências de Allais em uma extensão da teoria da decisão de Jeffrey – desta vez estendida a **proposições sobre chances**, isto é, proposições que descrevem distribuições de probabilidade objetivas. A ideia geral é que a conveniência de um determinado aumento ou diminuição na chance de algum resultado – por exemplo, no caso de Allais, um aumento de 0,01 na chance do resultado \$0 – pode depender de quais eram as chances antes do aumento ou diminuição. A extensão de Stefánsson e Bradley da teoria de Jeffrey para proposições sobre chances também é motivada pelo fato de que as teorias de decisão padrão não distinguem a aversão ao risco com respeito a algum bem das atitudes para com as quantidades desse bem. Tal aspecto é considerado problemático, por exemplo, por Hansson (1988), Rabin (2000) e Buchak (2013)²⁰.

²⁰ N.T.: Isto é, essas teorias não permitem distinguir (na respectiva função utilidade) a **aversão ao risco** propriamente dita (*grossa modo*, a ideia de \$50 pode ter maior valor que uma probabilidade de 0,5 de receber \$100) do **desconto marginal da utilidade** (*grossa modo*, que \$1 vale mais para alguém que tem \$50 do que para alguém que tem \$100).

5.2 Sobre a completude: crenças e desejos vagos

Conforme observado na Seção 4, as críticas ao requisito da UE de uma ordenação de preferências completa são motivadas por considerações epistêmicas e de desejo/valor. Quanto ao valor, muitos afirmam que um agente racional pode simplesmente encontrar duas opções *incomparáveis* devido às suas qualidades *incomensuráveis* (aqui, um uso proeminente desses termos será seguido, de acordo com o qual opções específicas podem ser descritas como incomparáveis em valor, enquanto propriedades gerais ou dimensões de valor podem ser descritas como incomensuráveis). Mais precisamente, as avaliações do agente sobre a conveniência de certas opções podem não ser representadas por nenhuma função de utilidade precisa. Da mesma forma, em relação à crença, alguns afirmam (notadamente, Joyce 2010 e Bradley 2017) que a evidência pode ser tal que não vincule um agente racional com graus precisos de crença mensuráveis por uma função de probabilidade única.

Existem várias representações alternativas e mais “difusas” (*fuzzier*) de desejo e crença que podem ser consideradas mais adequadas. Halpern (2003), por exemplo, investiga diferentes formas de conceituar e representar a incerteza epistêmica, uma vez que partimos das probabilidades. Presumivelmente, também existem várias maneiras de representar o desejo incerto. Aqui, o foco estará em apenas uma proposta que é popular entre os filósofos: o uso de *conjuntos* de funções de probabilidade e utilidade para representar a incerteza, respectivamente, na crença e no desejo. Esta é uma generalização mínima do modelo padrão da UE, no sentido de que as medidas de probabilidade e utilidade ainda estão presentes. Grosso modo, quanto mais severa a incerteza epistêmica, mais medidas de probabilidade sobre o espaço de possibilidades são necessárias para representar conjuntamente as crenças do agente. Essa noção de crença racional é conhecida como *probabilismo impreciso* (veja o verbete sobre probabilidades imprecisas na SEP). Da mesma forma, quanto mais severa a incerteza avaliativa, mais medidas de utilidade sobre o espaço de opções seguras são necessárias para representar conjuntamente os desejos do agente. Estritamente falando, não devemos tratar a crença e o desejo separadamente, mas, sim, falar das preferências incompletas do agente sendo representadas por um conjunto de pares de probabilidade e utilidade. Lembre-se do requisito de que preferências incompletas sejam *coerentemente extensíveis* (veja novamente a Seção 1); nesta representação, todos os pares probabilidade-utilidade equivalem a candidatos a extensões das preferências incompletas.

Surge então a questão: existe uma generalização conservadora da regra de decisão da UE que possa lidar com conjuntos de pares de probabilidade e utilidade? As regras de decisão do contendor são padronizadas em termos de funções de escolha que tomam como entrada algum conjunto de opções viáveis e retornam como saída um conjunto não vazio de escolhas admissíveis – que é um subconjunto das opções viáveis. Uma restrição básica dessas funções de escolha é que elas respeitam as preferências do agente nos casos em que as opções são de fato comparáveis. Isto é, se todos os pares de funções de probabilidade e utilidade que caracterizam as atitudes do agente concordam na classificação de duas opções, essas opções específicas devem ser classificadas de acordo. A restrição relevante sobre as funções de escolha é que "opções dominadas pela Utilidade Esperada" não são escolhas admissíveis, ou seja, se uma opção tem utilidade esperada menor do que outra opção, de acordo com todos os pares de funções de probabilidade e utilidade, então a primeira opção dominada não é uma escolha admissível. Observe que Levi (1986) tem uma condição um pouco mais restritiva de admissibilidade: se uma opção não tem UE máxima em pelo menos um par de funções de probabilidade e utilidade, então ela não é admissível. Em casos comuns onde os conjuntos de probabilidade e funções de utilidade são conjuntos convexos fechados, entretanto, a condição de Levi é equivalente à condição acima mencionada que exclui opções dominadas pela UE (SCHERVISH *et al.*, 2003).

O tratamento de opções genuinamente incomparáveis (aquelas que sobrevivem ao teste de admissibilidade apresentado acima, mas que ainda assim não são tais que o agente lhes seja indiferente) é onde começam as verdadeiras polêmicas. Veja Bradley (2017) para uma ampla discussão sobre as várias maneiras de proceder. Uma consideração para a qual frequentemente se apela para discriminar entre opções incomparáveis é a *cautela* (caution). A regra *Maxmin-UE*, por exemplo, recomenda escolher a ação com a maior utilidade mínima esperada (*vide* GILBOA; SCHMEIDLER, 1989; WALLEY, 1991). A regra é simples de usar, mas discutivelmente cautelosa demais, não prestando nenhuma atenção à amplitude da Utilidade Esperada. A regra α -*Maxmin*, por outro lado, recomenda executar a ação com a maior soma ponderada α das utilidades mínimas e máximas esperadas associadas a ela. Os pesos relativos para as utilidades mínimas e máximas esperadas podem ser considerados como refletindo o pessimismo do tomador de decisão em face da incerteza ou então seu grau de cautela (*vide* BINMORE, 2009).

Existem regras de escolha mais complicadas que dependem de uma representação mais rica da incerteza envolvendo uma noção de **confiança**. Por

exemplo, Klibanoff *et al.* (2005) propõem uma regra pela qual as escolhas são feitas entre opções incomparáveis com base na utilidade esperada ponderada pela confiança. Ele pressupõe que pesos podem ser atribuídos às várias utilidades esperadas associadas a um ato, refletindo a confiança do agente nos pares de probabilidade e utilidade correspondentes. Existem regras alternativas que apelam à confiança, mesmo na ausência de pesos cardinais precisos. Gärdenfors e Sahlin (1982), por exemplo, sugerem simplesmente excluir da consideração quaisquer funções de probabilidade (e utilidade) que caiam abaixo de um limite de confiança e, em seguida, aplicar a regra de *Maxmin-EU* com base no restante. A teoria da escolha de Hill (2013) é um tanto semelhante, embora os limites de confiança para pares de probabilidade e utilidade possam variar dependendo do problema de escolha (e o próprio termo “confiança” é usado de forma diferente). Existem outras propostas nas quais os atos são comparados em termos de quanta incerteza se pode tolerar (o que novamente depende dos níveis de confiança) e ainda assim ter uma opção satisfatória (*vide* BEN-HAIM, 2001). Essas regras são convincentes, mas levantam uma série de questões difíceis sobre como interpretar e medir as atitudes subjetivas extras que desempenham algum papel, como “quantidade de confiança em uma crença/desejo” e “nível satisfatório de desejabilidade”.

5.3 Desconhecimento

Recentemente, tem havido interesse em mais um desafio à teoria da utilidade esperada, a saber, o desafio do desconhecimento (*unawareness*)²¹. Na verdade, isso representa um desafio para todas as teorias normativas de escolha existentes. Para manter as coisas simples, devemos, no entanto, nos concentrar na teoria da utilidade esperada de Savage para ilustrar o desafio representado pelo desconhecimento.

²¹ N.T.: há poucas traduções para o português do termo *unawareness* neste contexto - em que um tomador de decisão não está ciente de alternativas/mundos possíveis relevantes. Opções como “falta de consciência” e “ignorância”, além de ambíguas, já denotam outros conceitos na literatura. Cabe alertar ao leitor, porém, que é possível que encontre traduções bastante distintas deste termo.

Como o leitor lembrará, Savage assume um conjunto de resultados possíveis **O** e outro conjunto de estados possíveis do mundo **S** e define o conjunto de atos, **F**, como o conjunto de todas as funções de **S** em **O**. Além disso, o teorema da representação tem sido interpretado como justificando a afirmação de que uma pessoa racional sempre realiza o ato em **F** que maximiza a utilidade esperada em relação a uma medida de probabilidade sobre **S** e uma medida de utilidade sobre **O**.

Agora, a teoria de Savage é neutra sobre como interpretar os estados **S** e os resultados em **O**. Por exemplo, a teoria é consistente com a interpretação de **S** e **O** como, respectivamente, os conjuntos de **todos** os estados e de resultados **logicamente possíveis**, mas também é consistente com a interpretação de **S** e **O** como, respectivamente, os conjuntos de estados e de resultados que algum criador de modelos/representação (“modelador”) reconhece, ou os conjuntos de estados e de resultados **que o próprio tomador de decisão reconhece**.

Se a teoria se destina a descrever o raciocínio de um tomador de decisão, as duas primeiras interpretações pareceriam inferiores à terceira. O problema com as duas primeiras interpretações é que o tomador de decisão pode **não estar ciente** de alguns dos estados e resultados logicamente possíveis, bem como alguns dos estados e resultados dos quais o modelador está ciente (podem-se identificar os estados e resultados dos quais o agente não tem ciência por referência àqueles dos quais o modelador está ciente).

Quando se trata de tomadores de decisão (parcialmente) não cientes, uma importante distinção pode ser feita entre, por um lado, o que podemos chamar de “desconhecimento do desconhecimento” – isto é, uma situação em que um tomador de decisão não percebe que pode haver algum resultado ou estado do qual não está ciente – e, por outro lado, “ciência do desconhecimento” – isto é, uma situação em que um tomador de decisão pelo menos suspeita que existe algum resultado ou estado do qual ele não tem conhecimento.

Do ponto de vista da tomada de decisão, o “desconhecimento do desconhecimento” não tem muito interesse. Afinal, se alguém nem mesmo está ciente da possibilidade de não estar ciente de algum estado ou resultado, então esse desconhecimento não pode desempenhar nenhum papel em nosso raciocínio sobre o que fazer. No entanto, modelos teóricos de decisão foram propostos sobre como uma pessoa racional responde à **diminuição** do desconhecimento (isso deve ser aplicado até mesmo a pessoas que anteriormente não estavam cientes de seu desconhecimento). Em particular, os economistas Karni e Vierø (2013, 2015) recentemente estenderam a condicionalização Bayesiana padrão para tais eventos

de aprendizagem. Sua teoria, o *Bayesianismo reverso*, afirma informalmente que o crescimento da ciência não deve afetar as razões de probabilidades dos estados/resultados de que o agente estava ciente antes do crescimento. Richard Bradley (2017) defende um princípio semelhante no contexto da estrutura mais geral ao estilo de Jeffrey, assim como o faz Roussos (2020); mas a visão é criticada por Steele e Stefánsson (2020, 2021) e por Mahtani (2021).

Em contraste, a consciência do desconhecimento parece ser de grande interesse do ponto de vista da tomada de decisão. Se você suspeita que existe algum estado possível, digamos, que você ainda não concebeu, e algum resultado correspondente, cujo conteúdo você desconhece, então você pode querer pelo menos chegar a uma opinião sobre quão provável você espera que esse estado seja, e quão bom ou ruim você espera que o resultado correspondente seja, antes de tomar uma decisão.

Várias pessoas sugeriram modelos para representar agentes que estão cientes de seu desconhecimento (*vide* WALKER; DIETZ, 2013; PIERMONT, 2017; KARNI; VIERØ, 2017). Steele e Stefánsson (2021) argumentam que pode não haver nada de especialmente distinto sobre como um tomador de decisão raciocina sobre estados/resultados dos quais está “ciente de que não está ciente” em termos da confiança que tem em seus julgamentos e de como gerencia o risco. Dito isso, vale a pena explorar mais a fundo a maneira como se chega a esses julgamentos de probabilidade e desejabilidade. Grant e Quiggin (2013a, 2013b), por exemplo, sugerem que esses julgamentos são feitos com base na indução de situações passadas em que se experimentou a diminuição do desconhecimento.

Em geral, a literatura sobre desconhecimento tem crescido rapidamente. Bradley (2017) e Steele e Stefánsson (2021) são novos tratamentos aprofundados desse tópico dentro da filosofia. Schipper mantém uma bibliografia sobre o assunto, principalmente com artigos em economia e ciência da computação, disponível em: <http://faculty.econ.ucdavis.edu/faculty/schipper/unaw.htm>. Acesso em: 12 jan. 2022.

6. Decisões sequenciais

As teorias de decisão de Savage e Jeffrey, bem como as de seus críticos, aparentemente dizem respeito a uma decisão única ou “de uma só vez”; foca-se a ordenação de preferência de um agente e, em última análise, sua escolha de ato, em relação a um determinado ponto no tempo. Alguém pode se referir a isso como

um problema de decisão **estático**. Surge então a questão de saber se este quadro é adequado para lidar com cenários mais complexos, em particular aqueles que envolvem uma série ou sequência de decisões; estes são chamados de problemas de decisão **sequencial**.

Pelo menos no papel, os modelos de decisão estáticos e sequenciais parecem muito diferentes. O modelo estático tem forma de tabela ou forma **normal**, em que cada linha representa um ato/opção disponível, e as colunas representam os diferentes estados possíveis do mundo que produzem um determinado resultado para cada ato. O modelo de decisão sequencial, por outro lado, possui forma de árvore ou **extensa** (como na Figura 1). Ele descreve uma série de pontos de escolha antecipados, onde os ramos que se estendem a partir de um ponto de escolha representam as opções naquele ponto de escolha. Algumas dessas ramificações levam a outros pontos de escolha, geralmente após a resolução de alguma incerteza devido a novas evidências.

- Essas diferenças básicas entre os modelos de decisão estático e sequencial levantam questões sobre como, de fato, eles se relacionam:
- Os modelos de decisão estáticos e sequenciais representam o mesmo tipo de problema de decisão? Em caso afirmativo, qual é a contraparte estática de um modelo de decisão sequencial?
- O cenário de decisão sequencial revela mais (des)vantagens da teoria da UE? De maneira mais geral, esse cenário esclarece as teorias normativas da escolha?

Essas questões acabam sendo bastante controversas. Elas serão abordadas uma a uma, depois que a cena for montada com a velha história sobre Ulisses.

6.1 Ulisses era racional?

Um problema de decisão sequencial bem conhecido é o que Ulisses enfrenta em sua jornada de volta à Ítaca no antigo épico de Homero. Ulisses deve fazer uma escolha sobre a maneira como navegará por uma ilha habitada por Sereias cantoras. Ele pode escolher navegar sem restrições ou amarrado ao mastro.

No primeiro caso, Ulisses mais tarde terá a escolha, ao ouvir as Sereias, de continuar navegando de volta para Ítaca ou de permanecer na ilha por tempo indeterminado. No último caso, ele não estará livre para fazer outras escolhas e o navio navegará para Ítaca passando pelas Sereias cantoras. O último resultado depende da sequência de escolhas que Ulisses faz. O problema de decisão de Ulisses é representado em forma de árvore (ou forma extensa) na Figura 1 (onde as duas caixas representam pontos de escolha para Ulisses).

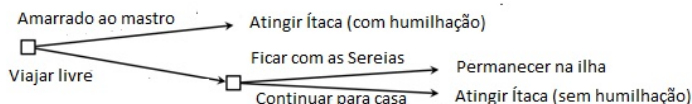


Figura 1. Problema de decisão de Ulisses

Somos informados de que, antes de embarcar, Ulisses preferia ouvir as Sereias livremente e então voltar para casa, em Ítaca. O problema é que Ulisses prediz que o seu eu do **futuro** não obedecerá a isto: se ele navegar livremente, será depois seduzido pelas Sereias e, em vez de continuar rumo à sua casa em Ítaca, permanecerá na ilha indefinidamente. Portanto, Ulisses raciocina que seria melhor ser amarrado ao mastro, porque é melhor a vergonha e o desconforto de ser amarrado ao mastro do que permanecer na ilha das Sereias para sempre.

É difícil negar que Ulisses fez uma escolha sábia ao ser amarrado ao mastro. Entretanto, alguns sustentam, que Ulisses não é um agente exemplar, já que o seu eu do presente deve jogar contra o seu eu do futuro, que será involuntariamente seduzido pelas Sereias. Embora Ulysses seja racional no **nó da primeira escolha** pelos padrões de decisão estáticos, podemos considerá-lo irracional **em geral** pelos padrões de decisão sequenciais, entendidos em termos do valor relativo das sequências de escolhas. A sequência de escolhas que Ulisses inevitavelmente busca é, afinal, sub-ótima. Teria sido melhor se ele pudesse navegar desamarrado e continuar para casa em Ítaca. Essa sequência poderia ter sido alcançada se Ulisses fosse continuamente racional **durante todo o tempo**; digamos, se em todos os momentos ele agisse como um maximizador da UE e mudasse suas crenças e desejos apenas de acordo com as normas Bayesianas (variantes da **condicionalização** padrão). Nessa leitura, os modelos de decisão sequencial introduzem considerações de “racionalidade-ao-longo-do-tempo”.

Embora a “racionalidade-ao-longo-do-tempo” possa ter importância na avaliação das preferências e das normas sobre como um agente altera essas preferências (pode-se ler a discussão na Seção 6.2 abaixo desta forma), permanece a importante questão de como um agente deve agir à luz de suas preferências em **qualquer momento**. Para tanto, o modelo de decisão sequencial pode ser visto com sucesso como uma ferramenta para ajudar a determinar a escolha racional em um determinado momento, assim como no modelo de decisão estático. A árvore de decisão sequencial é efetivamente uma forma de visualizar a série temporal de escolhas e eventos de aprendizagem que um agente **acredita** que enfrentará no futuro, dependendo de em qual parte da árvore de decisão encontrar-se-á. A questão-chave, então, é: como um agente deve escolher entre suas opções iniciais à luz de sua árvore de decisão projetada? Esta questão gerou uma quantidade surpreendente de controvérsia. Três abordagens principais para lidar com árvores de decisão sequenciais têm aparecido na literatura. São elas a abordagem **ingênua** ou **míope**, a abordagem **sofisticada** e a abordagem **resoluta**. Elas serão discutidas separadamente; será sugerido que as disputas podem não ser substanciais, mas antes indicam diferenças sutis na interpretação dos modelos de decisão sequenciais.

A chamada abordagem ingênua para tratar de decisões sequenciais serve como um contraste útil com as outras duas abordagens. O agente ingênuo assume que qualquer caminho através da árvore de decisão é possível, e então parte de qualquer caminho que seja considerado ótimo, dadas suas atitudes presentes. Por exemplo, um Ulisses ingênuo simplesmente presumiria que ele tem três estratégias gerais para escolher: ou ordenar à tripulação que o amarre ao mastro, ou não emitir tal ordem e depois permanecer na ilha das Sereias, ou não emitir tal ordem e, mais tarde, retomar seu curso. Ulisses prefere o resultado associado à última combinação e, portanto, inicia essa estratégia ao não ordenar que a tripulação o contenha. A Tabela 5 apresenta a contrapartida estática do problema de decisão do ingênuo Ulisses. Com efeito, esse modelo de decisão não leva em consideração o conhecimento atual de Ulisses de suas preferências futuras e, portanto, o aconselha a buscar uma opção que se prevê ser impossível.

Ação	Resultado
Ordenar ser amarrado ao mastro	Retornar a casa, alguma humilhação
navegar sem restrições e depois ficar com as Sereias	viver com as Sereias
navegar sem restrições e então voltar para Ítaca	chegar em casa, sem humilhação

Tabela 5. Problema de decisão “ingênuo” de Ulysses

Não há necessidade de insistir no fato de que a abordagem ingênua da escolha sequencial é apropriadamente nomeada. A marca registrada da abordagem sofisticada, em contraste, é sua ênfase no planejamento retroativo: o agente sofisticado não assume que todos os caminhos através da árvore de decisão (ou, em outras palavras, todas as combinações possíveis de escolhas nos vários nós de escolha) serão possíveis. O agente considera, ao contrário, o que estará inclinado a escolher em nós de escolha posteriores quando chegar à posição temporal em questão. O sofisticado Ulisses perceberia que, se chegar à ilha das Sereias desamarrado, vai querer parar ali indefinidamente, devido ao efeito transformador do canto das Sereias sobre suas preferências. Isso é então refletido na representação estática do problema de decisão, conforme a Tabela 6. Os estados aqui dizem respeito às preferências futuras de Ulisses, uma vez que ele chegue à ilha. Como o segundo estado tem (por suposição) probabilidade zero, os atos são decididos com base no primeiro estado, então Ulisses sabiamente escolhe ser amarrado ao mastro.

Ação	A seguir escolhe as Sereias ($p = 1$)	Depois escolhe Ítaca ($p = 0$)
Ordenar ser amarrado ao mastro	retornar a casa, alguma humilhação	retornar a casa, alguma humilhação
navegar sem restrições	viver com Sereias	retornar a casa sem humilhação

Tabela 6. Problema de decisão “sofisticado” de Ulisses

A escolha **resoluta** diverge da escolha sofisticada apenas sob certas condições que não são cumpridas por Ulisses, dada sua inexplicável mudança de atitudes. Os defensores da escolha resoluta normalmente defendem as teorias de decisão e preferências associadas que violam o axioma da Independência/Princípio da Coisa Segura (MCLENNEN, 1990; MACHINA, 1989). Confira, também, para uma discussão, Rabinowicz (1995) e Buchak (2013). Eles apelam à escolha resoluta para tornar essas preferências mais palatáveis no contexto de decisão sequencial (a ser discutido mais

detalhadamente na Seção 6.2 abaixo). De acordo com a escolha resoluta, em contextos apropriados, o agente deve em todos os pontos de escolha se ater à estratégia que foi inicialmente considerada melhor. A questão é se esse conselho faz sentido, dada a interpretação padrão de um modelo de decisão sequencial. O que significaria para um agente escolher contra suas preferências para cumprir um plano previamente selecionado? Isso parece desafiar a própria noção de preferência. Obviamente, um agente pode dar grande importância ao cumprimento de compromissos anteriores. Quaisquer preocupações com a integridade, entretanto, devem ser refletidas na especificação dos resultados e, portanto, nas preferências do agente no momento em questão. Isso é muito diferente de escolher em descompasso com suas preferências mais gerais, levando em consideração as situações relevantes em determinado momento.

Os defensores da escolha resoluta podem ter em mente uma interpretação diferente dos modelos de decisão sequencial, em que os “pontos de escolha” futuros não são realmente pontos nos quais um agente é livre para escolher de acordo com suas preferências no momento. Nesse caso, isso equivaleria a uma mudança sutil na questão ou problema de interesse. A seguir, a interpretação padrão dos modelos de decisão sequencial será assumida e, conseqüentemente, será assumido que os agentes racionais perseguem a abordagem sofisticada da escolha (LEVI, 1991; MAHER, 1992; SEIDENFELD, 1994).

6.2 Os axiomas da UE revisitados

Vimos que as árvores de decisão sequenciais podem ajudar um agente como Ulisses a fazer um balanço das consequências de sua escolha atual, para que ele possa refletir melhor sobre o que fazer **agora**. Entretanto, a literatura sobre escolha sequencial preocupa-se principalmente com questões mais ambiciosas. O cenário de decisão sequencial oferece efetivamente novas maneiras de “testar” teorias de preferência racional e normas para mudança de preferência (ou crença e desejo). A questão é se a teoria da decisão de um agente, neste sentido amplo, se mostra **dinamicamente inconsistente** ou autodestrutiva.

O argumento do *Dutch Book*²² diacrônico de Skyrms (1993) em favor da

²² N.T.: Dutch Book é uma aposta em que está garantido ao agente uma perda certa em razão da inconsistência de suas crenças ou preferências; v. a menção ao argumento da “bomba de dinheiro” na Seção 1.

condicionalização pode ser lido desta forma. Presume-se que o agente tenha preferências conforme a teoria da UE e adote uma abordagem sofisticada (com raciocínio retroativo) para problemas de decisões sequenciais. Skyrms mostra que qualquer agente que planeje aprender de uma maneira contrária à condicionalização fará escolhas contraproducentes em algumas situações de decisão sequencial especialmente planejadas. Em contraste, um agente que decide de acordo com a condicionalização nunca fará escolhas que sejam contraproducentes dessa maneira. O tipo de “escolhas autodestrutivas” em questão aqui são aquelas que geram uma perda certa. Ou seja, o agente escolhe uma estratégia que é certamente pior, segundo suas próprias preferências, do que outra estratégia que poderia ter escolhido, se sua regra de aprendizagem fosse tal que escolheria de forma diferente em um ou mais nós de decisão futuros.

Um argumento semelhante de “consistência dinâmica” pode ser usado para defender as preferências consistentes com a UE além do aprendizado de acordo com a condicionalização (*vide* HAMMOND, 1976, 1977, 1988b, 1988c). Como antes, presume-se que o agente adota uma abordagem sofisticada para problemas de decisão sequencial. Hammond mostra que apenas um agente totalmente bayesiano pode planejar seguir qualquer caminho em uma árvore de decisão sequencial que seja considerado ótimo a partir do nó de escolha inicial. Isso torna o agente bayesiano único, pois nunca fará “escolhas autodestrutivas” por conta de suas preferências e de suas normas para alterações de preferência. Nunca escolherá uma estratégia que é pior, a partir de seu ponto de vista, do que outra estratégia que poderia ter escolhido, se suas preferências fossem tais que teria escolhido de forma diferente em um ou mais nós de decisões futuras.

O argumento de Hammond para a teoria da UE, e a noção de consistência dinâmica que ele invoca, foi criticado em diferentes lugares, tanto por aqueles que defendem teorias que violam o axioma da Independência, mas mantêm os axiomas de Completude e Transitividade (ou seja, que defendem uma ordenação de preferências) da teoria da UE, quanto por aqueles que defendem teorias que violam Completude e Transitividade. Para uma discussão a esse respeito, confira Steele (2010). A abordagem adotada por alguns defensores das teorias que violam a Independência (MACHINA, 1989; MCCLENNEN, 1990) já foi mencionada: eles rejeitam a suposição de escolha sofisticada que sustenta os argumentos de consistência dinâmica. Seidenfeld (1988a, 1988b, 1994, 2000a, 2000b), ao invés disso, rejeita a noção de Hammond de consistência dinâmica em favor de uma noção mais sutil que discrimina entre as teorias que violam a Ordenação e aquelas

que violam apenas a Independência; os primeiros, ao contrário dos últimos, passam no teste de Seidenfeld, que se refere aos nós de decisões futuras em que o agente é indiferente entre as melhores opções. Este argumento também tem seus críticos (vide MCCLENNEN, 1988; HAMMOND, 1988a; RABINOWICZ, 2000).

Observe que os custos de qualquer desvio da teoria da UE são bem destacados por Al-Najjar e Weinstein (2009), em particular a possibilidade de aversão à informação gratuita e de aversão a oportunidades de mais escolhas no futuro. Kadane *et al.* (2008) e Bradley e Steele (2016) focam na perda certa que está associada ao pagamento para evitar evidências gratuitas. Mas, veja Buchak (2010, 2013) para uma discussão matizada desta questão em relação à racionalidade epistêmica versus a instrumental.

7. Observações finais

Vamos concluir resumindo as principais razões pelas quais a teoria da decisão, como descrita acima, é de interesse filosófico. Primeiro, a teoria da decisão normativa é claramente uma teoria (mínima) da racionalidade **prática**. O objetivo é caracterizar as atitudes de agentes que são praticamente racionais, e vários argumentos (estáticos e sequenciais) são tipicamente feitos para mostrar que certas “catástrofes” práticas acontecem a agentes que não satisfazem as restrições da teoria da decisão padrão. Em segundo lugar, muitas dessas restrições dizem respeito às **crenças** dos agentes. Em particular, a teoria da decisão normativa requer que os graus de crenças dos agentes satisfaçam os axiomas de probabilidade e que eles respondam a novas informações por condicionalização. Portanto, a teoria da decisão tem grandes implicações para os debates em epistemologia e filosofia da ciência; isto é, para teorias de racionalidade **epistêmica**.

Finalmente, a teoria da decisão deve ser de grande interesse para os filósofos da mente e da psicologia, e outros que estão interessados em como as pessoas podem compreender o comportamento e as intenções dos demais; e, de forma mais geral, como podemos interpretar o que se passa na mente de outras pessoas. Os teóricos da decisão normalmente assumem que o comportamento de uma pessoa pode ser totalmente explicado em termos de suas crenças e desejos. Mas, talvez mais interessante, alguns dos resultados mais importantes da teoria da decisão – os vários teoremas de representação, alguns dos quais discutidos aqui – sugerem que se uma pessoa satisfaz certos requisitos de racionalidade, então

podemos inferir suas crenças e desejos, e quão forte essas crenças e desejos são, a partir de suas disposições de escolha (ou preferências). Conforme discutido acima, o quanto esses teoremas realmente nos dizem é uma questão de debate. Mas, em uma leitura otimista desses resultados, eles nos asseguram que podemos falar de maneira significativa sobre o que se passa na mente de outras pessoas sem muito mais evidência além da informação sobre sua disposição para escolhas.

Referência bibliográfica

- AI-NAJJAR, N. I.; WEINSTEIN, J. The Ambiguity Aversion Literature: A Critical Assessment. **Economics and Philosophy**, v. 25, p. 249-284, 2009.
- ALLAIS, M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Américaine. **Econometrica**, v. 21, p. 503-546, 1953.
- ALT, F. Über die Meßbarkeit des Nutzens. **Zeitschrift für Nationalökonomie**, v. 7, p. 161-169, 1936.
- BEN-HAIM, Y. **Information-Gap Theory: Decisions Under Severe Uncertainty**. London: Academic Press, 2001.
- BERMÚDEZ, J. L. **Challenges to Decision Theory**. Oxford: Oxford University Press, 2009.
- BINMORE, K. **Rational Decisions**. Princeton: Princeton University Press, 2009.
- BOLKER, E. D. Functions Resembling Quotients of Measures. **Transactions of the American Mathematical Society**, v. 124, p. 292-312, 1966.
- BOLKER, E. D. A Simultaneous Axiomatisation of Utility and Subjective Probability. **Philosophy of Science**, v. 34, p. 333-340, 1967.
- BRADLEY, R. A Representation Theorem for a Decision Theory with Conditionals. **Synthese**, v. 116, p. 187-222, 1998.
- BRADLEY, R. Ramsey's Representation Theorem. **Dialectica**, v. 4, p. 484-497, 2004.
- BRADLEY, R. A Unified Bayesian Decision Theory. **Theory and Decision**, v. 63, p. 233-263, 2007.
- BRADLEY, R. **Decision Theory with a Human Face**. Cambridge: Cambridge University Press, 2017.
- BRADLEY, R.; STEFÁNSSON, H. O. Counterfactual Desirability. **British Journal for the Philosophy of Science**, v. 68, p. 485-533, 2017.

- BRADLEY, R.; STEFÁNSSON, H. O. Desire, Expectation and Invariance. **Mind**, v. 125, p. 691-725, 2016.
- BRADLEY, S.; STEELE, K. Can Free Evidence be Bad? Value of Information for the Imprecise Probabilist. **Philosophy of Science**, v. 83, p. 1-28, 2016.
- BROOME, J. **Weighing Goods**: Equality, Uncertainty and Time. Oxford: Blackwell, 1991a.
- BROOME, J. The Structure of Good: Decision Theory and Ethics. In: BACHARACH, M.; HURLEY, S. (ed.). **Foundations of Decision Theory**. Oxford: Blackwell, 1991b, p. 123-146.
- BROOME, J. Desire, Belief and Expectation. **Mind**, v. 100, p. 265-267, 1991c.
- BROOME, J. Can a Humean be moderate? In: FREY, G. R.; MORRIS, C. W. (ed.). **Value, Welfare and Morality**. Cambridge: Cambridge University Press, 1993, p. 51-73.
- BROWN, C. Consequentialize This. **Ethics**, v. 121, p. 749-771, 2011.
- BUCHAK, L. Instrumental rationality, epistemic rationality, and evidence-gathering. **Philosophical Perspectives**, v. 24, p. 85-120, 2010.
- BUCHAK, L. **Risk and Rationality**. Oxford: Oxford University Press, 2013.
- BUCHAK, L. Decision Theory. In: HITCHCOCK, C.; HÁJEK, A. (ed.). **Oxford Handbook of Probability and Philosophy**. Oxford: Oxford University Press, 2016, p. 789-814.
- BYRNE, A.; HÁJEK, A. David Hume, David Lewis, and Decision Theory. **Mind**, v. 106, p. 411-728, 1997.
- CHANG, R. The Possibility of Parity. **Ethics**, v. 112, p. 659-688, 2002.
- COLYVAN, M.; COX, D.; STEELE, K. Modelling the Moral Dimension of Decisions. **Noûs**, v. 44, p. 503-529, 2010.
- DAVIDSON, D.; MCKINSEY, J. C. C.; SUPPES, P. Outlines of a Formal Theory of Value, I. **Philosophy of Science**, v. 22, p. 140-160, 1955.
- DIETRICH, F.; LIST, C. A Reason-Based Theory of Rational Choice. **Noûs**, v. 47, p. 104-134, 2013.
- DIETRICH, F.; LIST, C. Reason-Based Choice and Context Dependence: An Explanatory Framework. **Economics and Philosophy**, v. 32, p. 175-229, 2016a.
- DIETRICH, F.; LIST, C. Mentalism Versus Behaviourism in Economics: a Philosophy-of-Science Perspective. **Economics and Philosophy**, v. 32, p. 249-281, 2016b.
- DIETRICH, F.; LIST, C. What Matters and How it Matters: A Choice-Theoretic Representation of Moral Theories. **The Philosophical Review**, v. 126, p. 421-479, 2017.

- DREIER, J. Rational Preference: Decision Theory as a Theory of Practical Rationality. **Theory and Decision**, v. 40, p. 249-276, 1996.
- DREIER, J. Structures of Normative Theories. **The Monist**, v. 76, p. 22-40, 1993.
- ELSTER, J.; ROEMER, J. E. (ed.). **Interpersonal Comparisons of Well-Being**. Cambridge: Cambridge University Press, 1993.
- ERIKSSON, L.; HÁJEK, A. What are Degrees of Belief. **Studia Logica**, v. 86, p. 183-213, 2007.
- GÄRDENFORS, P.; SAHLIN, N. E. Unreliability Probabilities, Risk Taking, and Decision Making. In: GÄRDENFORS, P.; SAHLIN, N. E. (ed.). **Decision, Probability and Utility**. Cambridge: Cambridge University Press, 1982, p. 313-334.
- GAIFMAN, H.; LIU, Y. A Simpler and more Realistic Subjective Decision Theory. **Synthese**, v. 195, p. 4205-4241, 2018.
- GILBOA, I.; SCHMEIDLER, D. Maxmin Expected Utility With Non-Unique Prior. **Journal of Mathematical Economics**, v. 18, p. 141-153, 1989.
- GOOD, I.J. On the Principle of Total Evidence. **British Journal for the Philosophy of Science**, v. 17, p. 319-321, 1967.
- GRANT, S.; QUIGGIN, J. Bounded Awareness, Heuristics and the Precautionary Principle. **Journal of Economic Behavior & Organization**, v. 93, p. 17-31, 2013a.
- GRANT, S.; QUIGGIN, J. Inductive Reasoning about Unawareness. **Economic Theory**, v. 54, p. 717-755, 2013b.
- GUALA, F. Has Game Theory Been Refuted? **Journal of Philosophy**, v. 103, p. 239-263, 2006.
- GUSTAFSSON, J. E. A Money-Pump for Acyclic Intransitive Preferences. **Dialectica**, v. 64, p. 251-257, 2010.
- GUSTAFSSON, J. E. The Irrelevance of the Diachronic Money-Pump Argument for Acyclicity. **The Journal of Philosophy**, v. 110, p. 460-464, 2013.
- HÁJEK, A.; PETTIT, P. Desire Beyond Belief. **Australasian Journal of Philosophy**, v. 82, p. 77-92, 2004.
- HALPERN, J. Y. **Reasoning About Uncertainty**. Cambridge: MIT Press, 2003.
- HAMMOND, P. J. Changing Tastes and Coherent Dynamic Choice. **The Review of Economic Studies**, v. 43, p. 159-173, 1976.
- HAMMOND, P. J. Dynamic Restrictions on Metastatic Choice. **Economica**, v. 44, p. 337-350, 1977.

- HAMMOND, P. J. Orderly Decision Theory: A Comment on Professor Seidenfeld. **Economics and Philosophy**, v. 4, p. 292-297, 1988a.
- HAMMOND, P. J. Consequentialism and the Independence Axiom. In: MUNIER, R. B. (ed.). **Risk, Decision and Rationality**. Dordrecht: D. Reidel, 1988b, p. 503-516.
- HAMMOND, P. J. Consequentialist Foundations for Expected Utility Theory. **Theory and Decision**, v. 25, p. 25-78, 1988c.
- HANSSON, B. Risk Aversion as a Problem of Conjoint Measurement. In: P. GÄRDENFORS, P.; SAHLIN, N. E. (ed.). **Decision, Probability and Utility**. Cambridge: Cambridge University Press, 1988, p. 136-158.
- HAUSMAN, D. M. Mistakes about Preferences in the Social Sciences. **Philosophy of the Social Sciences**, v. 41, p. 3-25, 2011a.
- HAUSMAN, D. M. **Preference, Value, Choice, and Welfare**. Cambridge: Cambridge University Press, 2011b.
- HEAP, S. H.; HOLLIS, M.; LYONS, B.; SUGDEN, R.; WEALE, A. **The Theory of Choice: A Critical Introduction**. Oxford: Blackwell Publishers, 1992.
- HILL, B. Confidence and Decision. **Games and Economic Behaviour**, v. 82, p. 675-692, 2013.
- JACKSON, F.; SMITH, M. Absolutist Moral Theories and Uncertainty. **The Journal of Philosophy**, v. 103, p. 267-283, 2006.
- JEFFREY, R. C. **The Logic of Decision**. New York: McGraw-Hill, 1965.
- JEFFREY, R. C. Preferences Among Preferences. **The Journal of Philosophy**, v. 71, p. 377-391, 1974.
- JEFFREY, R. C. Bayesianism With a Human Face. In: EARMAN, J. (ed.). **Testing Scientific Theories**. Minneapolis: University of Minnesota Press, 1983, p. 133-156.
- JOYCE, J. M. A Non-Pragmatic Vindication of Probabilism. **Philosophy of Science**, v. 65, p. 575-603, 1998.
- JOYCE, J. M. **The Foundations of Causal Decision Theory**. New York: Cambridge University Press, 1999.
- JOYCE, J. M. Levi on Causal Decision Theory and the Possibility of Predicting one's Own Actions. **Philosophical Studies**, v. 110, p. 69-102, 2002.
- JOYCE, J. M. A Defense of Imprecise Credences in Inference and Decision Making. **Philosophical Perspectives**, v. 24, p. 281-323, 2010.
- KADANE, J. B.; SCHERVISH, M. J.; SEIDENFELD, T. Is Ignorance Bliss? **The Journal of Philosophy**, v. 105, p. 5-36, 2008.

- KARNI, E.; VIERØ, M. L. Reverse Bayesianism: A Choice-Based Theory of Growing Awareness. **American Economic Review**, v. 103, p. 2790-2810, 2013.
- KARNI, E.; VIERØ, M. L. Probabilistic Sophistication and Reverse Bayesianism. **Journal of Risk and Uncertainty**, v. 50, p. 189-208, 2015.
- KARNI, E.; VIERØ, M. L. Awareness of Unawareness: A Theory of Decision Making in the Face of Ignorance. **Journal of Economic Theory**, v. 168, p. 301-325, 2017.
- KEENEY, R. L.; RAIFFA, H. **Decisions with Multiple Objectives**: Preferences and Value Tradeoffs. Cambridge: Cambridge University Press, 1993.
- KLIBANOFF, P.; MARINACCI, M.; MUKERJI, S. A Smooth Model of Decision Making Under Ambiguity. **Econometrica**, v. 73, p. 1849-1892, 2005.
- KNIGHT, F. **Risk, Uncertainty, and Profit**. Boston: Houghton Mifflin Company, 1921.
- KOOPMANS, T. C. Stationary Ordinal Utility and Impatience. **Econometrica**, v. 28, p. 287-309, 1960.
- KREPS, D. M. **Notes on the Theory of Choice**. Boulder: Westview Press, 1988.
- LAZAR, S. Deontological Decision Theory and Agent-Centred Options. **Ethics**, v. 127, p. 579-609, 2017.
- LEVI, I. **Hard Choices**: Decision Making Under Unresolved Conflict. Cambridge: Cambridge University Press, 1986.
- LEVI, I. Consequentialism and Sequential Choice. In: BACHARACH, M.; HURLEY, S. (ed.). **Foundations of Decision Theory**. Oxford: Basil Blackwell, 1991, p. 70-101.
- LEWIS, D. Desire as Belief. **Mind**, v. 97, p. 323-332, 1988.
- LEWIS, D. Desire as Belief II. **Mind**, v. 105, p. 303-313, 1996.
- LOOMES, G.; SUGDEN, R. Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. **The Economic Journal**, v. 92, p. 805-824, 1982.
- LOUISE, J. Relativity of Value and the Consequentialist Umbrella. **Philosophical Quarterly**, v. 4, p. 518-536, 2004.
- MACHINA, M. J. Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty. **Journal of Economic Literature**, v. 27, p. 1622-1668, 1989.
- MAHER, P. Diachronic Rationality. **Philosophy of Science**, v. 59, p. 120-141, 1992.
- MAHTANI, A. Awareness growth and dispositional attitudes. **Synthese**, 2020. DOI:10.1007/s11229-020-02611-5.
- MANDLER, M. A Difficult Choice in Preference Theory: Rationality Implies Completeness or Transitivity but Not Both. In: MILLGRAM, E. (ed.). **Varieties of Practical Reasoning**. Cambridge: MIT Press, 2001, p. 373-402, 2001.

- McCLENNEN, E. F. Ordering and Independence: A Comment on Professor Seidenfeld. **Economics and Philosophy**, v. 4, p. 298-308, 1988.
- McCLENNEN, E. F. **Rationality and Dynamic Choice**: Foundational Explorations. Cambridge: Cambridge University Press, 1990.
- MEACHAM, P.; CHRISTOPHER, J. G.; WEISBERG, J. Representation Theorems and the Foundations of Decision Theory. **Australasian Journal of Philosophy**, v. 89, p. 641-663, 2011
- PETERSON, M. **An Introduction to Decision Theory**. Cambridge: Cambridge University Press, 2009.
- PETTIT, P. Decision Theory and Folk Psychology. In: BACHARACH, M.; HURLEY, S. (ed.). **Foundations of Decision Theory**: Issues and Advances. Oxford: Blackwell, 1993, p. 147-175.
- PIERMONT, E. Introspective Unawareness and Observable Choice. **Games and Economic Behavior**, v. 106, p. 134-152, 2017.
- PORTMORE, D. W. Consequentializing Moral Theories. **Pacific Philosophical Quarterly**, v. 88, p. 39-73, 2007.
- RABIN, M. Risk Aversion and Expected-Utility Theory: A Calibration Theorem. **Econometrica**, v. 68, p. 1281-1292, 2000.
- RABINOWICZ, W. To Have one's Cake and Eat it, Too: Sequential Choice and Expected-Utility Violations. **Journal of Philosophy**, v. 92, p. 586-620, 1995.
- RABINOWICZ, W. Preference Stability and Substitution of Indifferents: A rejoinder to Seidenfeld. **Theory and Decision**, v. 48, p. 311-318, 2000.
- RABINOWICZ, W. Does Practical Deliberation Crowd out Self-Prediction? **Erkenntnis**, v. 57, p. 91-122, 2002.
- RAMSEY, F. P. Truth and Probability. In: BRAITHWAITE, R. B. (ed.). **The Foundations of Mathematics and other Logical Essays**. London: Trubner & Co., 1931, p. 156-198.
- RAMSEY, F. P. A Mathematical Theory of Saving. **The Economic Journal**, v. 38, p. 543-559, 1928.
- RAMSEY, F. P. Weight of the Value of Knowledge. **British Journal for the Philosophy of Science**, v. 41, p. 1-4, 1990.
- RESNIK, M. D. **Choices**: An Introduction to Decision Theory. Minneapolis: University of Minnesota Press, 1987.
- RINARD, S. No Exception for Belief. **Philosophy and Phenomenological Research**, v. 94, p. 121-143, 2017.

- ROUSSOS, J. **Policymaking Under Scientific Uncertainty**. London: London School of Economics and Political Science, 2020.
- SAVAGE, L. J. **The Foundations of Statistics**. New York: John Wiley and Sons, 1954.
- SCHERVISH, M. J.; SEIDENFELD, T.; KADANE, J. B.; LEVI, I. Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons. **Proceedings of the Third ISIPTA (JM)**, p. 496-510, 2003.
- SEIDENFELD, T. Decision Theory Without 'Independence' or Without 'Ordering'. **Economics and Philosophy**, v. 4, p. 309-315, 1988a.
- SEIDENFELD, T. Rejoinder [to Hammond and McClennen. **Economics and Philosophy**, v. 4, p. 309-315, 1988b.
- SEIDENFELD, T. When Normal and Extensive Form Decisions Differ. **Logic, Methodology and Philosophy of Science**, v. IX, p. 451-463, 1994.
- SEIDENFELD, T. Substitution of Indifferent Options at Choice Nodes and Admissibility: A Reply to Rabinowicz. **Theory and Decision**, v. 48, p. 305-310, 2000a.
- SEIDENFELD, T. The Independence Postulate, Hypothetical and Called-off Acts: A Further Reply to Rabinowicz. **Theory and Decision**, v. 48, p. 319-322, 2000b.
- SEN, A. Behaviour and the Concept of Preference. **Economica**, v. 40, p. 241-259, 1973.
- SEN, A. Rational Fools: A Critique of the Behavioural Foundations of Economic Theory. **Philosophy and Public Affairs**, v. 6, p. 317-344, 1977.
- SEN, A. Rights and Agency. **Philosophy and Public Affairs**, v. 11, p. 3-39, 1982.
- SKYRMS, B. A Mistake in Dynamic Coherence Arguments? **Philosophy of Science**, v. 60, p. 320-328, 1993.
- STALNAKER, R. C. **Inquiry**. Cambridge: MIT Press, 1987.
- STEELE, K. S. What are the Minimal Requirements of Rational Choice?: Arguments from the Sequential-Decision Setting. **Theory and Decision**, v. 68, p. 463-487, 2010.
- STEELE, K.; STEFÁNSSON, H. O. Belief Revision for Growing Awareness. **Mind**, jan., 2020. DOI: 10.1093/mind/fzaa056. 2020.
- STEELE, K.; STEFÁNSSON, H. O. **Beyond Uncertainty: Reasoning with Unknown Possibilities**. Cambridge University Press. 2021,
- STEFÁNSSON, H. O. Desires, Beliefs and Conditional Desirability. **Synthese**, v. 191, p. 4019-4035, 2014.
- STEFÁNSSON, H. O.; BRADLEY, R. What is Risk Aversion? **British Journal for the Philosophy of Science**, v. 70, p. 77-102, 2019.
- SUPPES, P. **Representation and Invariance of Scientific Structures**. Stanford: CSLI Publications, 2002.

- TEMKIN, L. **Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning**. Oxford: Oxford University Press, 2012.
- THOMA, J. Instrumental Rationality Without Separability. **Erkenntnis**, v. 85, p. 1219-1240, 2020a. DOI:10.1007/s10670-018-0074-9.
- THOMA, J. In Defense of Revealed Preference Theory. **Economics and Philosophy**, v. 1-25, 2020b. DOI:10.1017/s0266267120000073.
- TVERSKY, A. A Critique of Expected Utility Theory: Descriptive and Normative Considerations. **Erkenntnis**, v. 9, p.163-173, 1975.
- VILLEGAS, C. On Qualitative Probability $\setminus(\sigma)$ -Algebras. **Annals of Mathematical Statistics**, v. 35: p. 1787-1796, 1964.
- VALLENTYNE, P. Gimmicky Representations of Moral Theories. **Metaphilosophy**, v. 19, p. 253-263, 1988.
- VREDENBURGH, K. A Unificationist Defence of Revealed Preferences. **Economics and Philosophy**, v. 36, n. 1, p. 149-169, 2020. DOI:10.1017/s0266267118000524.
- VON NEUMANN, J.; MORGENSTERN, O. **Theory of Games and Economic Behavior**. Princeton: Princeton University Press, 1944.
- WALLEY, P. **Statistical Reasoning with Imprecise Probabilities**. New York: Chapman and Hall, 1991.
- WALKER, O.; DIETZ, S. A Representation Result for Choice Under Conscious Unawareness. Grantham Research Institute on Climate Change and the Environment. **Working Paper**, n. 59, 2011.
- ZYNDA, L. Representation Theorems and Realism about Degrees of Belief. **Philosophy of Science**, v. 67, n. 4, 2000.

Teoria dos Jogos e Ética*

Bruno Verbeek (Universidade de Leiden) e Christopher Morris (Universidade de Maryland)

Tradução: André Nascimento Pontes (UFAM)

Revisão: Mariana Kuhn de Oliveira (CEBRAP)

A teoria dos jogos é o estudo sistemático da escolha racional interdependente. Ela deve ser distinguida da teoria da decisão, que é o estudo sistemático da escolha individual (prática e epistêmica) em contextos paramétricos (isto é, em que o agente está escolhendo ou deliberando independentemente de outros agentes). A teoria da decisão tem várias aplicações em ética (DREIER, 2004; MELE; RAWLINGS, 2004).

A teoria dos jogos pode ser usada para explicar, prever e avaliar o comportamento humano em contextos em que o resultado da ação depende do que vários agentes escolhem fazer e onde suas escolhas dependem do que os outros escolhem fazer. (*vide verbete sobre a teoria dos jogos*). Consequentemente,

*VERBEEK, B.; MORRIS, C. *et al.* Game Theory and Ethics. In: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Winter Edition. Stanford, CA: The Metaphysics Research Lab, 2017. Disponível em: <https://plato.stanford.edu/archives/win2017/entries/game-ethics/>. Acesso em: 15 jan. 2022.

The following is the translation of the entry on Game Theory and Ethics de Bruno Verbeek, Christopher Morris, Keith Hankins and Peter Vanderschraaf, in the Stanford Encyclopedia of Philosophy. The translation follows the version of the entry in the SEP's archives at <<https://plato.stanford.edu/archives/win2020/entries/game-ethics/>>. This translated version may differ from the current version of the entry, which may have been updated since the time of this translation. The current version is located at <https://plato.stanford.edu/archives/win2017/entries/game-ethics/>. We'd like to thank the Editors of the Stanford Encyclopedia of Philosophy, mainly Prof. Dr. Edward Zalta, for granting permission to translate and to publish this entry.

a teoria dos jogos é relevante para a ética e é usada na filosofia moral e política de diversas maneiras.

Devemos nos concentrar na influência e uso da teoria dos jogos na ética e naquelas partes da teoria política que envolvem normas ou princípios de justiça, ignorando questões sobre instituições políticas e jurídicas e também questões sobre temas que lidam com virtudes morais.

É possível distinguir três diferentes tipos de investigações na literatura. A primeira, chamamos de *funcionalista*: a teoria dos jogos é usada para identificar a **função** da moralidade. Ela é usada para descrever o(s) problema(s) que ocorreria(m) na ausência da moralidade e desta descrição são extraídas as inferências sobre a função corretiva ou de melhoramento da moralidade. A segunda abordagem, o *contratarianismo*, usa a teoria dos jogos (especialmente a teoria da barganha) para formalizar a teoria do contrato social. Esta tradição mais antiga entende que as instituições ou normas políticas são justificadas na medida em que os agentes racionais concordariam com elas sob condições adequadas. A teoria da barganha tem sido usada para estabelecer, primeiro, que haverá acordo em tais condições, segundo, para prever o resultado deste processo de barganha e, finalmente, terceiro, a teoria dos jogos, especialmente a teoria dos jogos evolucionários, é usada para "recuperar" muitas normas ou práticas morais tradicionais. No que se segue, consideraremos cada uma dessas abordagens e os resultados e problemas que elas têm encontrado. Começaremos com alguns antecedentes históricos.

1. História

Em 1954 o filósofo britânico Richard Braithwaite deu uma palestra inaugural intitulada *Theory of Games as a Tool for the Moral Philosopher* [Teoria dos Jogos como uma Ferramenta para o Filósofo Moral] (BRAITHWAITE, 1955). Em sua fala, Braithwaite argumentou que muitas questões sobre justiça distributiva têm a mesma estrutura que "o problema da barganha". Este problema tinha sido analisado alguns anos antes por John Nash, o mais tarde ganhador do Prêmio Nobel, usando a teoria dos jogos (NASH, 1950). Braithwaite previu que a teoria dos jogos mudaria fundamentalmente a filosofia moral. Sua previsão veio menos de dez anos após a publicação de *Theory of Games and Economic Behaviour* [Teoria dos Jogos e Comportamento Econômico], de John von Neumann e Oskar Morgenstern, um livro que deu início a um ramo completamente novo das ciências sociais e da matemática

aplicada (VON NEUMANN; MORGENSTERN, 1944).

A introdução da teoria dos jogos na ética não foi um desenvolvimento inteiramente novo. As ideias da teoria dos jogos podem ser encontradas, por exemplo, nas obras de Thomas Hobbes e David Hume (GAUTHIER, 1969; KAVKA, 1986; HAMPTON, 1986; VANDERSCHRAAF, 1998). No entanto, a previsão de Braithwaite não se tornou realidade. A teoria dos jogos (ainda) não mudou fundamentalmente a ética. Dez anos após Braithwaite, Brian Barry publicou *Political Argument* [Discussão Política] e, alguns anos mais tarde, foi lançado *Convention* [Convenção], o trabalho seminal de David Lewis (BARRY, 1965; LEWIS, 1969). No final dos anos 60, apareceu a primeira de uma série de publicações de David Gauthier. Nelas, ele usou a teoria dos jogos para desenvolver sua teoria moral (GAUTHIER, 1967). Entretanto, até recentemente, a influência da teoria dos jogos na ética não tem sido tão grande em nenhum lugar como nas ciências sociais em geral. Apesar deste começo vacilante, a introdução da teoria dos jogos na filosofia moral produziu um fluxo crescente de publicações importantes.

2. Funcionalismo

A teoria dos jogos tem sido usada para analisar a **função** da moralidade. Um bom exemplo é *The Emergence of Norms* [A Emergência das Normas], de Edna Ullmann-Margalit, no qual ela argumenta que as normas morais permitem aos agentes cooperar e coordenar suas ações em situações em que a busca do interesse próprio os impediria (ULLMANN-MARGALIT, 1977). Seu exemplo agora clássico é o de dois artilheiros que enfrentam a escolha entre fugir do inimigo que está avançando ou ficar e operar suas armas. Suas armas estão localizadas em uma passagem estrategicamente importante. Se ambos ficarem, eles têm uma chance significativa de serem feridos, mas é certo que o avanço do inimigo será interrompido. Se ambos fugirem, o inimigo será capaz de tomar a passagem da montanha, ultrapassá-los e capturá-los. Se apenas um deles ficar enquanto o outro foge, o artilheiro corajoso morrerá em batalha, mas o outro artilheiro terá tempo suficiente para escapar em segurança. Supondo que ambos tentam sobreviver a esta provação, de preferência sem ferimentos, cada soldado tem motivos para fugir. A razão para isso é que eles estão envolvidos em um dilema do prisioneiro (*vide* Figura 1). Cada artilheiro tem a escolha entre fugir e permanecer e lutar. Esta escolha é representada nas linhas para o artilheiro #1 e nas colunas para o artilheiro #2. Cada célula da matriz representa o resultado de cada par de escolhas possíveis. Cada célula

tem um par de números. O número no canto inferior esquerdo de cada célula representa como o artilheiro #1 classifica este resultado, em relação aos outros resultados possíveis – representados por números de "utilidade". O número no canto superior direito representa a classificação deste resultado pelo artilheiro #2.

		Artilheiro #2	
		<i>Ficar</i>	<i>Fugir</i>
Artilheiro #1	<i>Ficar</i>	2 2	3 0
	<i>Fugir</i>	0 3	1 1

Figura 1

Considere o caso de #1. Suponha que #2 decida ficar e lutar. Nesse caso, para #1 é melhor fugir. Ele sobreviverá sem se ferir. Na representação formal da matriz, ele garantirá um ganho maior (3 em vez de 2). Suponha que #2 decida fugir. Novamente, #1 se sai melhor ao fugir. Ele sobreviverá à batalha, ainda que ele fique preso durante toda a guerra. Se ele ficasse e lutasse, certamente morreria; ao fugir ele garante uma posição mais alta (1 em vez de 0). O artilheiro #2 está na mesma posição que #1: para ele também, o que quer que o outro faça, ele se sai melhor ao fugir. Em resumo, cada artilheiro individualmente estaria melhor se fugisse, independentemente do que o outro faça. Contudo, continua sendo verdade – e para alguns, paradoxal – que ambos estariam em melhor situação se ambos mantivessem suas posições. O resultado da ação racional individual é Pareto-ineficiente (ou sub-ótimo).

Suponha que ambos compreendam a estrutura de sua difícil situação. Uma vez que eles veriam que cada um deles tem boas razões para fugir, poderiam tentar excluir essa possibilidade. Por exemplo, eles poderiam acorrentar um ao outro à arma, impedindo assim a fuga. Ullmann-Margalit argumentou que a situação dos artilheiros (isto é, o dilema do prisioneiro) é estruturalmente equivalente a muitas interações cotidianas regidas pela moralidade. Além disso, assim como o encadeamento mútuo compromete os artilheiros a permanecer e lutar, a moralidade compromete os agentes a evitar resultados Pareto-ineficientes ou sub-ótimos. A moralidade vincula os indivíduos às suas armas, por assim dizer. Nesta visão, a função da moralidade é evitar as falhas de racionalidade (MACKIE, 1977).

3. Problemas com o funcionalismo

Há vários problemas com esta análise funcional da moralidade. Primeiro, há alguns problemas bem conhecidos com explicações funcionalistas nas ciências sociais. O fato de que uma prática ou uma instituição tem uma função particular não precisa explicar nem seu surgimento nem sua manutenção. Pode-se argumentar, por exemplo, que a função do sistema público de educação é educar os jovens, que a função do Estado é servir aos interesses das classes dirigentes, ou que a função da religião é servir como ópio do povo. Entretanto, até que se possa demonstrar que estas aparentes funções são causalmente eficazes para a existência e manutenção do sistema educacional, do Estado ou da religião, respectivamente, nenhuma explicação foi dada. Da mesma forma, mesmo que normas e práticas morais sirvam para produção de resultados Pareto-superiores não realizáveis através de ações racionais individuais descoordenadas, nenhuma explicação da existência e persistência da moralidade é fornecida, a menos que se demonstre que esta função de alguma forma motiva a ação humana e ou de alguma outra forma é causalmente eficaz em produzir resultados mutuamente benéficos.

Em segundo lugar, pode-se questionar se a moralidade coincide com resultados mutuamente vantajosos ou Pareto-superiores da maneira sugerida. Muitos pensadores e pensadoras têm argumentado que muitas vezes somos moralmente obrigados a agir de maneiras que são desvantajosas para todos. Um exemplo óbvio é a proibição frequentemente afirmada contra a venda de si mesmo à escravidão. Pode muito bem ser vantajoso tanto para o escravo quanto para o patrão (o escravo poderia pagar suas dívidas e o patrão teria uma solução prática para o trabalho doméstico diário), mas isso é moralmente e legalmente proibido.

Terceiro, a abordagem funcionalista assume claramente que as exigências da moralidade entram em conflito com a racionalidade individual. A moralidade deve corrigir problemas da preocupante ineficiência de Pareto, que seria o resultado de uma ação racional individual irrestrita (interdependente). Na abordagem funcionalista, o agente moral parece *ipso facto* ser irracional (salvo considerações de expiação de culpa e arrependimento). Isto torna a questão "por que ser moral?" uma petição de princípio. O funcionalismo inviabiliza uma resposta a esta questão.

E, finalmente, em quarto lugar, o objetivo das abordagens funcionalistas é de interesse limitado aos teóricos morais. O funcionalismo parece buscar

explicações para o surgimento e persistência de normas e práticas morais. Os teóricos e teóricas morais não estão interessados principalmente em tais explicações. Ao contrário, eles normalmente procuram compreender a moralidade com o objetivo de determinar o que devemos fazer ou o que somos obrigados a fazer. É a moralidade enquanto guia para a ação e para a vida que é o principal interesse do filósofo e da filósofa moral. A moralidade aqui é **normativa**, uma fonte de orientação. Suponha que houvesse uma explicação funcional plausível de normas morais particulares. Essa explicação mostra que sou, de fato, **obrigado** a seguir essas normas quando elas se aplicam a mim? Parece haver uma diferença entre (a) determinar a(s) função(ões) da moralidade e (b) verificar se um determinado conjunto de normas e práticas são, de fato, as que devemos seguir. Não está claro como esta pergunta é respondida pelas abordagens funcionalistas.

4. Teoria da barganha e contratarianismo

Como vimos acima, uma das críticas ao funcionalismo é que ele não explica a conexão entre a escolha individual e o surgimento e persistência de normas morais. A moralidade é introduzida como algo fora da escolha racional individual. Em resposta a esta dificuldade, muitos teóricos e teóricas têm tentado entender a moralidade como resultado da escolha racional individual. Grosso modo, podemos distinguir duas estratégias. Primeiro, há aqueles que modelam a moralidade como o resultado de uma escolha única de uma grande coleção de agentes, a comunidade moral. Segundo, há aqueles que abordam a moralidade como resultado de uma série de interações repetidas em pequena escala. Discutiremos esta segunda abordagem na seção 7. Aqui discutiremos a abordagem que considera a moralidade como o resultado pretendido das interações entre os agentes racionais sob circunstâncias igualmente ideais. Esta é uma ideia antiga em filosofia moral e política: é a ideia do **contrato social** (*vide* verbete sobre o **contratarianismo**). A moralidade é interpretada como o resultado de um processo de barganha.

A introdução à teoria dos jogos, especialmente aquelas partes da teoria que se preocupam com a barganha (a chamada teoria dos jogos cooperativos e a teoria da barganha), tem estimulado o interesse pela teoria do contrato social nas últimas décadas. John Harsanyi, Richard Braithwaite, John Rawls, Brian Barry e David Gauthier usaram a teoria dos jogos e da decisão para formular versões da teoria (HARSANYI, 1955; BRAITHWAITE, 1955; BARRY, 1965; RAWLS, 1971;

GAUTHIER, 1986). Invocando a teoria da barganha, eles tentaram mostrar (1) que agentes racionais em uma situação de barganha adequadamente idealizada concordarão com uma distribuição específica e única dos benefícios da cooperação, (2) como se parece essa distribuição, (3) que essa distribuição determina o que é justo e (4), no caso de Gauthier, que os agentes racionais cumprirão os termos da barganha.

É importante para essas teorias como exatamente a situação de barganha é caracterizada. Gauthier, assim como muitos outros, pensa nela como um dilema do prisioneiro. Isto é, a situação das partes na posição de barganha ideal é estruturalmente equivalente à situação dos artilheiros, tal como a descrevemos acima. Sem qualquer cooperação, os artilheiros estão condenados a fugir e passar o restante da guerra em cativeiro. Suponhamos que seja possível fazer acordos vinculativos nesta situação. Isso resolve o problema preocupante da eficiência de Pareto? Não, porque não é óbvio **como** os benefícios da cooperação serão distribuídos. Pode parecer que neste caso só há uma maneira de distribuí-los, mas as aparências enganam. Os artilheiros poderiam decidir seguir uma **estratégia mista**. Uma estratégia mista é uma loteria sobre as estratégias disponíveis de cada indivíduo. Por exemplo, os artilheiros poderiam decidir fugir com uma probabilidade de, digamos, $1/3$ e ficar e lutar com uma probabilidade de $2/3$. (Deve-se notar que a ideia de uma estratégia mista geralmente é introduzida no contexto das chamadas *utilidades cardinais*. Enquanto antes os números na matriz (0, 1, 2 e 3) significavam apenas a **classificação** do resultado, aqui se assume que os números fornecem algumas informações sobre a **classificação relativa** do resultado. Por exemplo, a utilidade de "2" do resultado cooperativo significa que o agente é indiferente entre este resultado e uma aposta que oferece a ele "0" (o pior resultado) com probabilidade de $1/3$ e "3" (seu melhor resultado) com probabilidade de $2/3$. (Para uma discussão detalhada da teoria da utilidade cardinal, ver a seção 3.5 do verbete sobre interpretações de probabilidade). Daqui em diante, assumiremos que os números na matriz são utilidades cardinais).

Os artilheiros percebem que cada um deles pode obter individualmente o único e pior resultado da não cooperação. Isto significa que o resultado do acordo entre eles deve ser pelo menos tão bom quanto o resultado não cooperativo. Portanto, a distribuição com a qual eles concordarão deve ser pelo menos 1. Suponha que os artilheiros tenham um par de dados. Agora eles podem realizar distribuições cooperativas que não sejam 2 cada. Por exemplo, se eles concordarem em lançar ambos os dados e se um total de 6 ou menos aparecer, o artilheiro #1 fugirá (chegando assim a um valor de utilidade de 3). Entretanto, se o total de ambos os

dados for superior a 6, o artilheiro #1 ficará e lutará contra o inimigo (chegando assim ao seu pior resultado de 0). A utilidade esperada deste acordo para o artilheiro #1 é de $\frac{5}{12} \cdot 3 + \frac{7}{12} \cdot 0 = 1,25$, enquanto o artilheiro #2 pode esperar 1,75 deste acordo. Desta forma, os artilheiros podem obter toda uma gama de resultados, variando as chances que melhoram o resultado não cooperativo. Estes resultados formam a área de barganha (vide Figura 2).

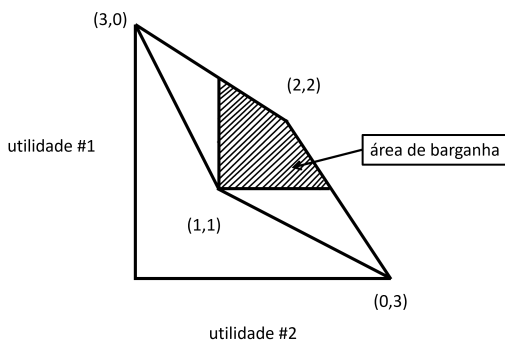


Figura 2

Intuitivamente pode parecer simples que o resultado do acordo entre os artilheiros #1 e o #2 será (2,2). Formalmente, isto é tudo menos simples. Cada resultado que dá a cada artilheiro uma utilidade esperada de mais de 1 parece racionalmente aceitável. Qual deles será selecionado pelos artilheiros racionais? Dentro da teoria da barganha, a parte da teoria dos jogos que trata destes problemas, há duas abordagens que procuram responder a esta pergunta (BINMORE, 1998, cap. 1). Primeiro, há a **abordagem axiomática** tradicional, tal como desenvolvida no contexto da teoria dos jogos cooperativos. Este ramo da teoria dos jogos assume que, uma vez que os agentes racionais tenham chegado a um acordo, eles o respeitarão. A tarefa do teórico e da teórica é considerar a área de barganha e determinar qual(is) resultado(s) satisfaria(m) uma série de requisitos razoáveis de um resultado racional das negociações. Coisas como os nomes das partes envolvidas não devem ser importantes para o resultado, ao passo que suas preferências sim. Esta abordagem tem sido muito influente na teoria do contrato social que faz uso da teoria dos jogos. Harsanyi, Rawls, Barry e Gauthier se utilizaram de abordagens axiomáticas para justificar sua versão favorita. Seus vereditos no caso dos artilheiros

é o mesmo: o racional é concordar com uma distribuição que dá a cada artilheiro uma utilidade esperada de 2.

A abordagem axiomática não dá atenção à estrutura do processo de negociação. Todos os dados iniciais que ela requer como input são informações sobre as recompensas das partes. Embora seja verdade que às vezes não importa exatamente como o processo de negociação é estruturado, às vezes ele é muito importante. Por exemplo, se é o caso que #1 pode fazer uma reivindicação e tudo o que #2 pode fazer é aceitar ou recusar, #1 tem sua melhor opção oferecendo a #2 uma utilidade esperada de 1,00001 e reivindicando para si uma utilidade de 2,99999. Dadas as regras do processo de negociação, #2 terá que aceitar a proposta, já que a alternativa é (ligeiramente) pior. Por outro lado, se as regras permitem a troca de reivindicações e ofertas, a situação é bastante diferente. Portanto, se você quiser prever qual será o resultado do processo de negociação entre agentes racionais, é crucial conhecer em detalhes as regras de negociação, assim como a área de barganha. Além disso, é importante saber se as partes cumprirão o acordo. Se este não for o caso, é improvável que as partes envolvidas aceitem tal acordo em vez de um acordo que será vinculativo.

Portanto, é melhor pensar no processo de barganha como uma série de movimentos possíveis em um jogo que precede o jogo que os artilheiros enfrentam. Esta é a segunda abordagem, que considera os processos de barganha como **jogos não cooperativos**. A solução para tal jogo corresponde então à solução do processo de barganha. Nesta abordagem, é preciso prestar muita atenção aos detalhes. Consequentemente, a análise é complicada e muitas vezes confusa (esta é outra razão pela qual a abordagem axiomática é tão atraente para alguns).

Entretanto, é muito bem possível que a solução ao jogo e a solução baseada na abordagem axiomática sejam idênticas. De fato, isto é o que seria de se esperar se a solução axiomática proposta fosse de todo plausível. Esta intuição é a força motriz do chamado *programa de Nash* (NASH, 1950). Este programa visa a avaliar soluções axiomáticas ao verificar se o resultado de um jogo de negociação leva ao mesmo resultado. O sucesso do programa de Nash é crucial para a plausibilidade das teorias axiomáticas clássicas do contrato social. Tais teorias consideram a moralidade como o resultado de negociações (hipotéticas) entre agentes idealmente racionais, mas não se preocupam em explicitar exatamente **como** as partes chegam a este resultado. Consequentemente, se não houver pelo menos a promessa de tal análise detalhada, como é prometido pelo programa de Nash, o resultado que elas apresentam carece de plausibilidade. Para tratamentos mais recentes do problema da barganha, confira os trabalhos de Rubinstein (1982) e Binmore (1998).

5. Moralidade por acordo

Uma das teorias contratarianistas mais influentes em circulação é a de David Gauthier. No entanto, sua teoria é diferente de outras abordagens contratarianistas, não apenas em seu uso extensivo da teoria dos jogos e da barganha, mas também no seguinte aspecto. Uma das dificuldades que sinalizamos com relação à abordagem funcionalista é que ela não dá resposta à pergunta "por que ser moral?". É aqui que a teoria contratarianista de Gauthier se distingue das de Rawls, Harsanyi e outros. Gauthier não só usa a teoria da barganha para determinar, como Rawls e Harsanyi procuraram fazer, o conteúdo dos princípios morais fundamentais; ele também tenta mostrar que os agentes racionais agirão moralmente. Por esta razão, discutiremos essa teoria em maior detalhe.

A teoria moral de Gauthier, a "moralidade por acordo" (GAUTHIER, 1986), é uma teoria sobre a natureza e a racionalidade da moralidade. (*vide* a seção 3 do verbete sobre **contratarianismo**). Ela consiste em quatro partes. A primeira é uma interpretação da razão prática e da condição natural da humanidade, em grande parte familiar aos teóricos da escolha racional e aos teóricos contratarianistas da moral (GAUTHIER, 1986, cap. 2-4). Em seguida, há uma exposição dos princípios de conduta que os agentes racionais hipoteticamente concordariam; um tipo de "contrato social" (GAUTHIER, 1986, cap. 5). A terceira parte é uma interpretação revisionista controversa da racionalidade prática essencial ao seu argumento que busca mostrar que praticamente todos, em circunstâncias normais, têm motivos para aceitar e respeitar as limitações impostas por estes princípios (GAUTHIER, 1986, cap. 6). Finalmente, Gauthier argumenta que os princípios em questão são princípios de moralidade, um argumento que faz referência implícita a um relato funcionalista de normas morais (GAUTHIER, 1986, cap. 7-8). A terceira parte é a resposta de Gauthier à pergunta "por que ser moral?". Ela toca em algumas questões muito fundamentais na teoria dos jogos e da decisão, e é por isso que a discutimos um pouco mais aqui.

Como Hobbes já percebeu, uma coisa é chegar a um acordo; outra é cumprir sua parte em um acordo. A moralidade, pelo menos como é tradicionalmente concebida, muitas vezes exige que sacrifiquemos nossos interesses ou objetivos. Isto é, pelo menos à primeira vista, o contrário do que a racionalidade requer. A resposta de Gauthier a isto é argumentar que concebemos erroneamente a

racionalidade prática, e até mesmo a racionalidade instrumental, se pensarmos que o objetivo da racionalidade é determinar de qualquer forma direta a maneira pela qual devemos raciocinar ou deliberar. O **objetivo** da racionalidade – fazer o melhor possível – não determina necessariamente nosso **princípio de decisão**; por exemplo, escolher a melhor alternativa em cada momento de escolha. Nos termos da concepção de racionalidade fundamentada na maximização da utilidade que ele aceitou até recentemente (GAUTHIER, 2008²³), Gauthier argumenta que o objetivo de maximizar a utilidade não significa que devemos, em cada ponto de decisão, maximizar a utilidade. Ao invés disso, devemos raciocinar de forma a maximizar a utilidade. Assim como às vezes ocorre de nos darmos melhor ou pelo menos bem não tendo buscado esses resultados, pode também às vezes ocorrer de o curso de ação para maximizar a utilidade não ser maximizar a utilidade em cada ponto de decisão. Dado que nosso modo de raciocínio ou deliberação em si afeta nossas perspectivas, nossos objetivos ou propósitos às vezes são mais bem servidos quando não procuramos fazer o melhor em cada ponto de decisão.

A discussão de Gauthier em *Morals by Agreement* [Moralidade por Acordo] é conduzida em termos de "disposições para escolher" e, especificamente, de "maximização restrita"; a disposição de cooperar com outros cooperadores mesmo em circunstâncias em que desertar é mais vantajoso. Em trabalho posterior, Gauthier desenvolve sua abordagem revisionista da racionalidade prática em termos de planos e intenções racionais e de modos de deliberação. Se concedermos que os agentes podem fazer melhor em qualquer número de circunstâncias agindo de maneiras que não sejam "diretamente maximizadoras", o problema passa a ser o de determinar como pode ser racional agir como um maximizador restrito. No livro, Gauthier assume que se nossas disposições para escolher são racionais, então nossas escolhas determinadas por estas disposições também são racionais. Um grande número de especialistas tem seguido Thomas Schelling ao argumentar que muitas vezes é racional fazer coisas que são irracionais, mas esses teóricos reivindicam também que essas coisas não deixam de ser irracionais mesmo nas circunstâncias específicas. Gauthier pensa que se um curso de ação é melhor do que qualquer outro em seus efeitos, então, sob certas condições, pode ser racional

²³ N.T.: O texto original, na versão utilizada para a tradução, indicava que essa obra estava no prelo. Como a obra, atualmente, já se encontra disponível, optou-se por indicar o ano de sua publicação.

adotá-lo e ter a intenção de desempenhar suas partes mesmo que algumas delas não sejam, da posição ocupada no momento da execução, a melhor coisa a fazer em termos dos objetivos ou propósitos do indivíduo. Ele procura, portanto, estabelecer que se um modo de deliberação ou um plano de ação é racional, então agir de acordo com ele pode ser racional, mesmo que tal ação exija fazer coisas que não são consideradas ótimas da posição ocupada no momento da ação. A ação baseada em princípios restringe a ação de alguém; e é racional ser assim restrita. Assim, se Gauthier estiver certo, pode ser racional obedecer a certas normas ou princípios mesmo quando eles exigem agir de maneiras que não são as melhores da perspectiva ocupada no momento da ação. Grande parte do trabalho de Gauthier desde *Morals by Agreement* desenvolve e defende esta abordagem revisionista de racionalidade prática. (vide GAUTHIER, 1994, 1996, 1998a, 1998b). Para uma abordagem revisionista alternativa, confira McClennen (1990).

A defesa de Gauthier da "maximização restrita" constitui uma revisão importante das teorias dos jogos e da decisão ortodoxas. A teoria ortodoxa enfoca a racionalidade das ações no momento da escolha. O próprio modo de deliberação sobre ações está fora do escopo da teoria (ou ainda, a teoria ortodoxa se apresenta como tal modo de deliberação.) Alguns críticos têm argumentado contra a inclusão do modo de deliberação no escopo da teoria (VELLEMAN, 1997). A maioria dos teóricos e teóricas dos jogos, entretanto, argumentam que se for viável escolher o modo de deliberação, esta escolha pode, ela própria, ser modelada como um movimento em um jogo de decisão mais complexo, incluindo assim a proposta de Gauthier na teoria ortodoxa (BINMORE, 1994, p. 179-182).

6. Alguns problemas com a abordagem contratarianista

A abordagem contratarianista, e a teoria de Gauthier não é diferente neste aspecto, pressupõe uma conexão fundamental entre racionalidade e moralidade, assim como o funcionalismo. Entretanto, diferentemente do projeto funcionalista, a abordagem contratarianista apresenta um argumento sofisticado sobre a razão pela qual isso deveria ser assim. Normas morais (ou instituições, ou qualquer que seja o objeto da teoria em questão) só são racionalmente aceitáveis de acordo com a tradição contratarianista se não houver um arranjo alternativo viável em que todas as partes envolvidas estariam em melhor situação. Podemos tornar esta afirmação mais vívida. Imagine que as partes estão barganhando sobre qual norma usar para

compartilhar um bolo. As partes racionais não concordariam com uma norma que deixaria algum bolo em cima da mesa para ser desperdiçado. Da mesma forma com normas morais: os agentes racionais não concordariam com uma norma que poderia deixar as vantagens mútuas inexploradas. Portanto, de acordo com a concepção de moralidade dos contratarianistas, é necessariamente o caso de que a moralidade correta leva a resultados Pareto-eficientes. Por esta razão, o contratarianismo de escolha racional é frequentemente considerado como revisionista em suas implicações. A alegação não é de que o bom senso ou a moralidade comum conduzem a resultados Pareto-eficientes (se seguidos). Em vez disso, a alegação do contratarianismo de escolha racional é que a interpretação correta das normas morais vinculantes é aquela que implica que se essas normas forem seguidas, os resultados serão Pareto-eficientes.

Críticos há muito argumentam que não está claro por que o resultado de um hipotético acordo deveria influenciar o que os agentes fora das circunstâncias idealizadas do "contrato social" deveriam fazer. Alguns têm argumentado que contratos hipotéticos (ou promessas) não vinculam. No entanto, essa é uma compreensão errônea da natureza destas teorias; o acordo racional hipotético não se destina a ser como uma promessa. Ao contrário, o acordo é antes de tudo heurístico, um mecanismo projetado para determinar a natureza e o conteúdo de princípios justos e mutuamente benéficos.

Enquanto as observações acima abordam todas as formas de contratarianismo, há alguns problemas específicos com versões que dependem em larga medida da teoria da barganha desenvolvida pela teoria dos jogos, como é o caso da teoria de Gauthier entre outras. O mais fundamental parece ser a plausibilidade do programa de Nash: existe realmente uma solução racional para todos os problemas de barganha que pode ser especificada e testada com o uso da teoria dos jogos não cooperativos? Da mesma forma, como podemos ter certeza de que há sempre uma solução única? Ou, estariam os problemas de barganha, até certo ponto, subdeterminados? A pluralidade de conceitos de solução de barganha que são discutidos na teoria da barganha é uma indicação de problema nesse caso. Há razões para duvidar que a abordagem da barganha pela teoria dos jogos possa realmente auxiliar na previsão do resultado das negociações de agentes racionais. Tanto a abordagem axiomática quanto a abordagem dos jogos não cooperativos partem do pressuposto de que existe um resultado único e racional de tais negociações. Embora isso possa ser plausível em algumas situações, está longe de ser óbvio que este seja sempre o caso. Ou seja, o resultado das negociações frequentemente

parece racionalmente subdeterminado (SUGDEN, 1991). Fatores não racionais, tais como saliência, precedência, etc., são muito mais importantes para determinar o resultado de tais negociações do que a teoria ortodoxa dos jogos e da barganha nos leva a acreditar.

7. Teoria dos jogos evolucionários e ética

Há também outro tipo de preocupação que naturalmente leva ao terceiro grande movimento da teoria dos jogos e da ética. Os contratarianistas, como Gauthier, entendem as normas fundamentais que nos governam como derivadas de uma situação de escolha (hipotética) que teria um número muito grande de agentes barganhando sobre diferentes princípios ou arranjos sociais. Entretanto, é uma questão em aberto se essa é uma maneira apropriada de modelar o processo de escolha racional que leva ao surgimento da moralidade.

Neste ponto, há uma diferença fundamental com a terceira maneira pela qual a teoria dos jogos tem sido aplicada à ética. Esta terceira maneira é a teoria dos jogos evolucionários. Em vez de considerar a moralidade como o resultado pretendido de um complexo processo de barganha em larga escala entre agentes plenamente informados e plenamente racionais, a abordagem evolucionária afasta-se de todas estas suposições. Em primeiro lugar, a moralidade é vista como o efeito colateral não intencional das interações dos agentes. Em segundo lugar, a moralidade emerge de uma série de interações repetidas entre pequenos grupos de agentes (a maioria dos modelos lida apenas com interações de duas pessoas). Para colocar isto em termos funcionalistas: a moralidade não é para resolver um problema, mas problemas recorrentes. Em terceiro lugar, em vez de supor informação completa e racionalidade plena, a teoria dos jogos evolucionária faz suposições menos exigentes das habilidades cognitivas e deliberativas dos agentes. Isto pode levar a resultados fundamentalmente diferentes.

Podemos ilustrar isto da seguinte forma. Rousseau descreve o estado da natureza como sendo parecido com a Caça ao Veado {Stag Hunt} (ROUSSEAU, 1964, p. 166-167). Para um tratamento contemporâneo deste jogo, confira Skyrms (2004). Imagine dois caçadores que podem escolher caçar lebres. Suas chances de caçar uma lebre não são afetadas pelas ações de outros. Entretanto, ambos preferem comer carne de veado no jantar, mas se forem caçar veado, só serão bem-sucedidos em conjunto.

		#2	
		<i>veado</i>	<i>lebre</i>
#1	<i>veado</i>	3 3	2 0
	<i>lebre</i>	0 2	2 2

Figura 3: Caça ao Veado

Suponhamos que #1 e #2 coordenem suas ações em (lebre, lebre). Este equilíbrio é muito Pareto-inferior a (veado, veado). Enquanto a escolha contratarianista resultaria em que (veado, veado) é a norma correta a ser estabelecida, a teoria dos jogos evolucionários nos ensina que é improvável que o equilíbrio Pareto-eficiente seja selecionado em um processo de repetidas interações. Além disso, o equilíbrio Pareto-eficiente é instável: desvios ocasionais deste equilíbrio levarão a população como um todo a coordenar suas ações para obter (lebre, lebre) em vez de (veado, veado).

Presumivelmente, isso é verdade para algumas de nossas normas reais – sejam elas sociais, legais ou morais. Elas podem ser deficientes em relação a outras normas; especialmente aquelas que se originam de situações de escolha social idealizada da teoria moral contratarianista. No entanto, a maioria de nossas normas reais são frequentemente estáveis, e não está claro que tenhamos motivos para nos afastarmos delas. Portanto, somos deixados com a dúvida sobre se as normas descobertas pela teoria da barganha conforme desenvolvida no âmbito da teoria dos jogos são viáveis para a maioria das sociedades, comunidades e grupos. Como "dever" implica "poder", temos razões para duvidar que a abordagem contratarianista nos dê uma descrição correta da moralidade que devemos seguir.

O principal resultado da abordagem evolucionária até agora é a "recuperação" de muitas intuições e normas morais existentes. Assim, teóricos e teóricas da teoria dos jogos evolucionários, ao escrever sobre ética (assim como os filósofos e filósofas morais que utilizam da teoria dos jogos evolucionários), têm mostrado que entre os agentes não tão racionais podem surgir muitas das normas de coordenação e cooperação que são objeto de investigação das teorias morais mais tradicionais

(SUGDEN, 1986; BINMORE, 1994, 1998; SKYRMS, 1996). Além disso, Skyrms (1996) e outros demonstraram que agentes autointeressados, diferentemente, desenvolverão heurísticas de raciocínio como a *Regra de Ouro* (faça aos outros o que você quer que seja feito para você) e uma versão da "maximização restrita" de Gauthier sob circunstâncias apropriadas. Ou seja, eles mostram que a evolução favorece não apenas o surgimento de modelos de comportamento que se conformam aos padrões morais, mas também favorece o desenvolvimento de heurísticas cognitivas que têm todas as características do raciocínio moral.

8. Algumas observações sobre a abordagem evolucionária

A maioria dos autores e autoras que adotaram a abordagem evolucionária apontam prontamente que esta abordagem evita muitas críticas oferecidas às duas abordagens anteriores. Primeiro, a abordagem evolucionária fornece uma explicação genuína do surgimento e persistência de normas morais. As normas são o efeito colateral não intencional das ações de agentes (limitadamente) racionais e emergem no processo de interações repetidas. Na abordagem evolucionária, a "função" de uma norma moral é selecionar um equilíbrio estável em uma situação em que há mais de uma possibilidade. Desta forma, normas estáveis podem ser Pareto-ineficientes. Não existe um elo fundamental entre eficiência e moralidade na abordagem evolucionária. Seu foco é o equilíbrio e não a eficiência. Esta é também a razão pela qual um agente em tal população deve seguir essa norma. Ou seja, o fato de que os outros membros de uma população seguem uma norma explica por que, e justifica que, um indivíduo em tal população também o faça. Como consequência, a abordagem evolucionária fornece uma resposta à pergunta "por que ser moral?". Seguir uma norma existente é individualmente racional. Além disso, nenhuma revisão não ortodoxa da teoria da escolha precisa ser aceita para alcançar este resultado, o que é uma grande vantagem sobre as reivindicações de Gauthier sobre a "maximização restrita".

No entanto, há também alguns motivos para desconfiar do sucesso da abordagem evolucionária, pois, assim como a abordagem funcionalista e ao contrário do projeto contratarianista, seu objetivo é oferecer explicação. A teoria dos jogos evolucionários é usada principalmente para explicar o surgimento e a estabilidade das normas existentes. Ela não fornece os instrumentos para se criticar o conteúdo dessas normas. Ela não fornece nenhuma justificativa para um código de conduta

que seja decididamente moral, entretanto, confira os trabalhos de Binmore (1994, 1998).

Esta tendência é especialmente preocupante quando vemos na literatura sobre a evolução do comportamento explicações para disposições desagradáveis, como a propensão dos homens ao estupro, a inclinação humana para fazer distinções de status baseadas em critérios como o de raça, entre outros. Portanto, não está claro até que ponto esta abordagem oferece uma alternativa às teorias morais existentes. Ela é provavelmente melhor entendida como uma forma de teoria social, ainda que ambivalente quanto ao seu status de teoria empiricamente informada ou de forma de teorização *a priori* (SUGDEN, 2001). É claro, pode-se pensar que a teoria dos jogos evolucionários não é uma alternativa à teoria moral, mas sim um veículo para minar ou desmascarar reivindicações morais. Se a fonte de nossas disposições e julgamentos morais é essencialmente a mesma que as inclinações desagradáveis mencionadas acima, então talvez devêssemos concluir que nossos julgamentos morais são falsos ou injustificados e que nossas disposições morais não são confiáveis. A teoria dos jogos evolucionários, nesta interpretação, apoiaria uma espécie de ceticismo moral (*vide* a seção 1 do verbete **ceticismo moral**). Algumas respostas a este ceticismo podem ser encontradas, por exemplo, em Gibbard (1990).

9. Algumas aplicações abstratas da teoria dos jogos na ética

Independentemente dos méritos das três abordagens que discutimos acima, há algumas ideias extraordinárias que a aplicação da teoria dos jogos oferece ao teórico e à teórica moral. Como observamos acima, há muitos jogos com múltiplos equilíbrios. Este é especialmente o caso de rodadas iteradas de jogos particulares, como o dilema do prisioneiro. Uma das implicações deste fato é que, na medida em que estes jogos são representações ou modelos úteis de nossas interações sociais, temos razões para esperar grande indeterminação no mundo. Como consequência, temos motivos para desconfiar de teóricos e teóricas morais que reivindicam universalidade e generalidade para suas recomendações normativas específicas (HARDIN, 1988, 2003).

Em segundo lugar, a teoria dos jogos deixa claro que em qualquer população suficientemente grande podemos esperar determinadas combinações de disposições comportamentais. Considere o já bastante popular Jogo do Falcão e do Pombo (SMITH, 1982):

		agente #2	
		Pombo	Falcão
agente #1	Pombo	2 2	3 1
	Falcão	1 3	0 0

Figura 4: O jogo do Falcão e do Pombo

Os dois equilíbrios em estratégias puras no jogo 2 2 simples resultam da adoção de uma estratégia diferente por cada jogador. Se pensarmos nas estratégias de "Falcão" e "Pombo" como representando disposições de caráter morais, então podemos ter razões para esperar que as populações humanas consistirão em agentes com diferentes tipos de caráter, por assim dizer (*vide* FRANK, 1988; SMITH, 1982; SKYRMS, 1996). Além disso, dada esta análise, está longe de ser claro que o teórico e a teórica moral estejam em qualquer posição de recomendar a mesma disposição, isto é, a mesma virtude, para todos os agentes desta população: alguns deveriam ser "Falcões" (*Hawk*), outros "Pombos" (*dove*), confira, também, Kuhn (2004).

Enquanto as duas últimas observações apontam para ideias originais para a teoria moral, não podemos evitar mencionar algumas das críticas que foram formuladas contra a aplicação da teoria dos jogos à ética. As mais fundamentais dizem respeito à antropologia implícita do agente racional. A questão é se tudo que é relevante sobre o agente para a teoria moral pode ser capturado pela imagem unidimensional do homem racional proposta pela teoria dos jogos. Supõe-se que o agente seja completamente caracterizado por sua ordenação de preferência sobre os resultados e por suas crenças em cada etapa do jogo. Entretanto, distinções moralmente importantes – por exemplo, entre diferenças de caráter – não têm lugar nesta caracterização.

Podemos ilustrar essa preocupação com a forma como o conceito de reputação é utilizado nos modelos de cooperação altruísta. A teoria dos jogos em seu estágio mais atual fez uso da noção de **reputação** de um jogador a fim de explicar a cooperação em rodadas iteradas de jogos como o dilema do prisioneiro

(KREPS; WILSON, 1982). Ao repetir-se muitas vezes as rodadas do dilema do prisioneiro, compensa ter uma reputação de cooperação. No entanto, não está claro o que significa exatamente ter uma **reputação** nesses contextos. Normalmente, uma reputação é o que geralmente se acredita sobre o **caráter** de uma pessoa. Nesses modelos, por outro lado, uma reputação é simplesmente um histórico dos **movimentos** do jogador em jogos semelhantes. Há uma diferença moralmente relevante entre essas duas concepções de reputação. Em que acreditamos quando aprendemos que um comerciante é honesto? Normalmente supomos que isso significa que ele é o tipo de pessoa que não vai enganar outras pessoas, por exemplo, os clientes, mesmo em situações em que isso lhe valha a pena. Por que o comerciante faria isso? Enquanto outro comerciante não trapaceia porque (ou quando) não vale a pena, nosso comerciante é honesto e não trapaceia por causa de sua honestidade, ou seja, de seu caráter. Normalmente, isto faz uma grande diferença na forma como julgá-los estes dois comerciantes. Ambos se comportam cooperativamente, mas somente o último é digno de louvor por sua honestidade. A teoria dos jogos e a teoria da utilidade geralmente não abrem espaço para esta distinção (*vide* MORRIS, 1999). Nesse ponto, é também relevante Brennan e Pettit (2004).

10. Conclusão

A maioria dos autores e autoras contemporâneos em ética que utilizam a teoria dos jogos em seus trabalhos são ou contratarianistas ou teóricos evolucionários. As duas abordagens representam duas combinações diferentes de teoria dos jogos e ética. A tradição contratarianista, com sua ênfase em agentes plenamente racionais e barganha, representa um uso mais tradicional da teoria dos jogos. A abordagem evolucionária, por outro lado, com sua ênfase em agentes racionais limitados e interações repetidas, é mais recente. Para a maioria dos especialistas na área, uma síntese destas abordagens parece altamente desejável, sendo os trabalhos de Binmore (1994, 1998), até hoje, a única tentativa.

Referência bibliográfica

BARRY, B. **Political Argument**, International library of philosophy and scientific method. London: Routledge & Kegan Paul, 1965.

- BINMORE, K. **Playing Fair**. Cambridge: The MIT Press. 1994. (Game Theory and the Social Contract; v. 1).
- BINMORE, K. **Just Playing**. Cambridge: The MIT Press, 1998. (Game Theory and the Social Contract, v. 2).
- BINMORE, K. **Natural Justice**. New York: Oxford University Press, 2005.
- BRAITHWAITE, R. B. **Theory of games as a tool for the moral philosopher**. Cambridge: Cambridge University Press, 1955.
- BRENNAN, G.; PETTIT, P. **The Economy of Esteem: An Essay on Civil and Political Society**. Oxford: Oxford University Press, 2004.
- DREIER, J. Decision Theory and Morality. In: **The Oxford Handbook of Rationality**. Oxford: Oxford University Press, 2004.
- FRANK, R. **Passions within Reason**. London: W. W. Norton & Company, Inc, 1988.
- GAUTHIER, D. Morality and Advantage. **Philosophical Review**, v. 76, p. 460-475, 1967.
- GAUTHIER, D. **The Logic of 'Leviathan': the Moral and Political Theory of Thomas Hobbes**. Oxford: Clarendon Press, 1969.
- GAUTHIER, D. **Morals by Agreement**. Oxford: Clarendon Press, 1986.
- GAUTHIER, D. Assure and Threaten. **Ethics**, v. 104, p. 690-721, 1994.
- GAUTHIER, D. Commitment and Choice: An Essay on the Rationality of Plans. In: FARINA, F.; HAHN, F.; VANNUCCI, S. **Ethics, Rationality, and Economic Behaviour**. Oxford: Oxford University Press, 1996, p. 217-244.
- GAUTHIER, D. Intention and Deliberation. In: **Modeling Rationality, Morality, and Evolution**. DANIELSON, P. (ed.). Oxford: Oxford University Press, 1998a, p. 41-54.
- GAUTHIER, D. Rethinking the Toxin Puzzle. In: COLEMAN, J.; MORRIS, C. **Rational Commitment and Social Justice: Essays for Gregory Kavka**. Cambridge: Cambridge University Press, 1998b, p. 47-58.
- GAUTHIER, D. Friends, Reasons and Morals. In: VERBEEK, B. **Reasons and Intentions**. Aldershot: Ashgate, 2008, p.17-36
- GIBBARD, A. **Wise Choices, Apt Feelings: A Theory of Normative Judgment**. Cambridge: Harvard University Press, 1990.
- HAMPTON, J. **Hobbes and the Social Contract Tradition**. Cambridge: Cambridge University Press, 1986.
- HARDIN, R. Bargaining For Justice. **Social Philosophy and Policy**, v. 5, p. 65-74, 1988.
- HARDIN, R. **Indeterminacy and Society**. Princeton: Princeton University Press. 2003.
- HARSANYI, J. C. Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility. **Journal of Political Economy**, v. 63, p. 309-321, 1955.

- KAVKA, G. **Hobbesian Moral and Political Theory**. Princeton: Princeton University Press, 1986.
- KREPS, D. M.; WILSON, R. Reputation and Imperfect Information. **Journal of Economic Theory**, v. 27, n. 2, p. 253-279, 1982.
- KUHN, S. T. Reflections on Ethics and Game Theory. **Synthese**, v. 141, n.1, p. 1-44, 2004.
- LEWIS, D. **Convention: A Philosophical Study**. Cambridge: Harvard University Press, 1969.
- MACKIE, J. **Ethics**. London: Penguin Books Ltd, 1977.
- MCCLENNEN, E. F. **Rationality and Dynamic Choice: Foundational Explorations**. Cambridge: Cambridge University Press, 1990.
- MELE, A.; RAWLINGS, P. (ed.). **The Oxford Handbook of Rationality**. Oxford: Oxford University Press, 2004.
- MORRIS, C. W. What is this Thing Called 'Reputation'? **Business Ethics Quarterly**, v. 9, n. 1, p. 87-102, 1999.
- NASH, J. The Bargaining Problem. **Econometrica**, v. 18, n. 2, p. 155-162, 1950.
- RAWLS, J. **A Theory of Justice**. Cambridge: Harvard University Press, 1971.
- ROUSSEAU, J. J. **Discours sur l'origine et les fondements de l'inégalité parmi les hommes**. Paris: Éditions Gallimard, 1964. (Oeuvres complètes, v. III).
- RUBINSTEIN, A. Perfect Equilibrium in a Bargaining Model. **Econometrica**, v. 50, n. 1, p. 97-109, 1982.
- SKYRMS, B. **Evolution of the Social Contract**. Cambridge: Cambridge University Press, 1996.
- SKYRMS, B. **The Stag Hunt and the Evolution of Social Structure**. Cambridge: Cambridge University Press, 2004.
- SMITH, J. M. **Evolution and the Theory of Games**. Cambridge: Cambridge University Press, 1982.
- SUGDEN, R. **The Economics of Rights, Co-operation and Welfare**. Oxford: Basil Blackwell, 1986.
- SUGDEN, R. Rational Bargaining. In: BACHARACH, M.; HURLEY, S. **Foundations of Decision Theory**. Oxford: Basil Blackwell, 1991.
- SUGDEN, R. The Evolutionary Turn in Game Theory. **Journal of Economic Methodology**, v. 8, n. 1, p. 113-130, 2001.
- ULLMANN-MARGALIT, E. **The Emergence of Norms**. Oxford: Oxford University Press, 1977.
- VANDERSCHRAAF, P. The Informal Game Theory in Hume's Account of Convention. **Economics and Philosophy**, v. 14, n. 2, p. 215-247, 1998.

- VELLEMAN, J. D. Deciding how to Decide. *In*: CULITTY, G.; GAUT, B. **Ethics and Practical Reason**. Oxford: Oxford University Press. 1997, p. 29-52.
- VERBEEK, B. **Instrumental Rationality and Moral Philosophy**: an essay on the virtues of cooperation. Dordrecht: Kluwer Academic Publishers, 2002.
- VERBEEK, B. The Authority of Norms. **American Philosophical Quarterly**, v. 44, n. 3, p. 245-258, 2007.
- VERBEEK, B. Conventions and Moral Norms: the legacy of Lewis. **Topoi** 27, v. 1-2, p. 73-86, 2008.
- VON NEUMANN, J.; MORGENSTERN, O. **Theory of Games and Economic Behavior**. Princeton: Princeton University Press, 1944.

Teoria da Escolha Social*

Christian List (LMU)**

Tradução: Mariana Kuhn de Oliveira (CEBRAP)

Revisão: Thomas Hyeono Kang (ESPM/ FGV- São Paulo)

A teoria da escolha social é o estudo de processos e procedimentos de decisão coletiva. Não é uma teoria única, mas um conjunto de modelos e resultados que concernem a agregação de dados individuais (por exemplo, votos, preferências, julgamentos, bem-estar) em resultados coletivos (por exemplo, decisões coletivas, preferências, julgamentos, bem-estar). As questões centrais são: como pode um grupo de indivíduos escolher um resultado vencedor (por exemplo, política pública, candidatura eleitoral) a partir de um dado conjunto de opções? Quais são as propriedades de diferentes sistemas de votação? Quando um sistema de votação

*LIST, C. Social Choice Theory. In: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Winter Edition. Stanford, CA: The Metaphysics Research Lab, 2013. Disponível em: <https://plato.stanford.edu/archives/win2013/entries/social-choice/>. Acesso em: 15 jan. 2022.

The following is the translation of the entry on Social Choice Theory de Christian List, in the Stanford Encyclopedia of Philosophy. The translation follows the version of the entry in the SEP's archives at <<https://plato.stanford.edu/archives/win2013/entries/social-choice/>>. This translated version may differ from the current version of the entry, which may have been updated since the time of this translation. The current version is located at <https://plato.stanford.edu/entries/social-choice/> We'd like to thank the Editors of the Stanford Encyclopedia of Philosophy, mainly Prof. Dr. Edward Zalta, for granting permission to translate and to publish this entry.

**Agradeço aos editores, aos revisores, Franz Dietrich, Iain McLean e Michael Morreau pelos comentários.

é democrático? Como pode uma coletividade (por exemplo, eleitorado, legislatura, tribunais colegiados, painel de especialistas ou comitê) alcançar preferências ou julgamentos coletivos coerentes em alguns assuntos a partir das preferências ou julgamentos de seus membros individuais? Como podemos ordenar diferentes alternativas sociais em uma ordem de bem-estar social? Os teóricos da escolha social estudam essas questões não apenas olhando para exemplos, mas também desenvolvendo modelos gerais e provando teoremas.

Tendo como seus pioneiros, no século XVIII, Nicolas de Condorcet e Jean-Charles de Borda e, no século XIX, Charles Dodgson (também conhecido como Lewis Carrol), a teoria da escolha social decolou no século XX com os trabalhos de Kenneth Arrow, Amartya Sen e Duncan Black. Sua influência se estende pelos campos da economia, da ciência política, da filosofia, da matemática e, recentemente, da computação científica e da biologia. Além de contribuir para nossa compreensão dos procedimentos de decisão coletiva, a teoria da escolha social tem aplicação nas áreas de desenho institucional, economia do bem-estar e epistemologia social.

1 História da teoria da escolha social

1.1 Condorcet

Os dois estudiosos mais frequentemente associados ao desenvolvimento da teoria da escolha social são o francês Nicolas de Condorcet (1743–1794) e o americano Kenneth Arrow (nascido em 1921)²⁴. Condorcet foi um pensador liberal na era da Revolução Francesa que foi perseguido pelas autoridades revolucionárias por as ter criticado. Após um período se escondendo, ele acabou sendo preso, porém aparentemente não foi imediatamente identificado, e morreu na prisão (para mais detalhes sobre Condorcet, veja McLean e Hewitt, 1994). Em seu *Ensaio sobre a Aplicação de Análise à Probabilidade da Decisões Majoritárias* [*Essai sur l'Application de l'Analyse*] (1785), ele defendeu um sistema particular de votação, votação majoritária em pares, e apresentou suas duas ideias mais proeminentes. A primeira,

²⁴ N.T.: Kenneth Arrow faleceu em 2017, após a última atualização da versão em língua inglesa desse artigo.

conhecida como o *teorema do júri de Condorcet*, afirma que se cada membro de um júri tem uma chance independente e igual melhor que o acaso, mas pior que a perfeição, de julgar corretamente se um réu é culpado (ou sobre alguma outra proposição factual), a maioria dos jurados tem a maior probabilidade de estar correta do que cada jurado individual, e a probabilidade de um julgamento majoritário correto se aproxima de 1 à medida que o tamanho do júri aumenta. Assim, sob certas condições, a regra da maioria é boa em “rastrear a verdade” (e.g., Grofman, Owen e Feld 1983; List e Goodin, 2001).

A segunda ideia de Condorcet, normalmente chamada de *paradoxo de Condorcet*, é a observação de que as preferências da maioria podem ser “irracionais” (especificamente, intransitivas) mesmo quando as preferências individuais são “racionais” (especificamente, transitivas). Suponha, por exemplo, que um terço de um grupo prefere alternativa x a y e y a z , e o último terço prefere z a x e x a y . Há maiorias (de dois terços) para x contra y , para y contra z e para z contra x : um “ciclo”, o que viola a transitividade. Além disso, nenhuma alternativa é uma *vencedora de Condorcet*, uma alternativa que ganhe de, ou ao menos empate com, cada uma das outras alternativas em competições majoritárias em pares.

Condorcet antecipou um tema chave na teoria da escolha social: a regra da maioria é ao mesmo tempo o método plausível de se realizar decisões coletivas e também sujeito a alguns problemas surpreendentes. Resolver ou contornar esses problemas continua sendo uma das preocupações centrais da teoria da escolha social.

1.2 Arrow e sua influência

Enquanto Condorcet investigou um método particular de votação (votação majoritária), Arrow, que ganhou o Prêmio Nobel em Economia em 1972, introduziu uma abordagem **geral** ao estudo da agregação de preferências, parcialmente inspirado por seu professor de lógica, Alfred Tarski (1901-1983), com quem ele havia aprendido teoria das relações enquanto era um estudante de graduação na *City College of New York* (SUPPES, 2005). Arrow considerou um grupo de métodos de agregação *possíveis*, os quais ele denominou *funções de bem-estar social*, e perguntou quais deles satisfaziam certos axiomas ou desideratos. Ele provou que, surpreendentemente, não existe método para agregação de preferências de dois ou mais indivíduos sobre três ou mais alternativas em preferências coletivas, onde este método satisfaz cinco axiomas aparentemente plausíveis discutidos abaixo.

Esse resultado, conhecido como *teorema da impossibilidade de Arrow*, provocou muito trabalho e muitos debates em teoria da escolha social e economia do bem-estar. William Riker (1920-1993), que inspirou a escola Rochester de ciência política, interpretou o teorema como prova matemática da impossibilidade de uma democracia populista (RIKER, 1982). Outros, mais proeminentemente Amartya Sen (nascido em 1933), que ganhou o Prêmio Nobel de 1998, buscaram mostrar que as preferências ordinais são insuficientes para fazer escolhas sociais satisfatórias. Comentaristas também questionaram se os desideratos de Arrow para um método de agregação são tão inócuos quanto alegados ou se eles deveriam ser mais brandos.

As lições do teorema de Arrow dependem, em parte, de como interpretamos uma função social de bem-estar arrowiana. O uso de preferências ordinais como “agregandos” pode ser mais facilmente justificado se interpretarmos a regra de agregação como um *método de votação* do que se a interpretarmos como um *método de avaliação de bem-estar*. Sen argumentou que, quando um planejador social busca ordenar diferentes alternativas sociais em uma ordem de bem-estar social (empregando, assim, alguma regra de agregação como um método de avaliação de bem-estar), pode ser justificável usar informações adicionais para além das preferências ordinais, como medidas de bem-estar interpessoalmente comparáveis (SEN, 1982).

O próprio Arrow afirmava que a comparação interpessoal de utilidades não possui significado e [...] que não há significado relevante para comparações de bem-estar na mensurabilidade de utilidade individual. (1951/1963, p. 9)

Esta visão foi influenciada pela economia neoclássica, associada a teóricos como Vilfredo Pareto (1848-1923), Lionel Robbins (1898-1984), John Hicks (1904-1989), vencedor do Prêmio Nobel de Economia com Arrow, e Paul Samuelson (1915-2009), outro ganhador do Nobel. O teorema de Arrow demonstra as implicações drásticas dos pressupostos “ordinalistas” do pensamento neoclássico.

Hoje em dia, a maioria dos teóricos da escolha social superaram as interpretações negativas iniciais do teorema de Arrow e estão interessados nas compensações (*trade-offs*) envolvidas na descoberta de procedimentos de decisão satisfatórios. Sen promoveu esta interpretação “possibilista” da teoria da escolha social (por exemplo, em sua palestra na premiação do Nobel de 1998).

Nessa abordagem, o método axiomático de Arrow é talvez até mesmo mais influente do que o teorema da impossibilidade. Sobre o método axiomático, confira Thomson (2000). O tipo de resultado paradigmático em trabalho axiomático contemporâneo é o “teorema da caracterização”. Aqui o objetivo é identificar um

conjunto de condições plausíveis necessárias e suficientes que caracterizam de maneira única uma solução particular (ou classe de soluções) para um determinado tipo de problema de decisão coletiva. Um exemplo inicial é a caracterização de Kenneth May (1952) da regra da maioria, discutida abaixo.

1.3 Borda, Carroll, Black e outros

Condorcet e Arrow não são as únicas figuras fundantes da teoria da escolha social. Contemporâneo e compatriota de Condorcet, Jean-Charles de Borda (1733-1799) defendeu um sistema de votação que é frequentemente visto como alternativa proeminente à votação majoritária. A *contagem de Borda*, formalmente definida posteriormente, evita o paradoxo de Condorcet, violando, entretanto, uma das condições de Arrow, a da *independência de alternativas irrelevantes*. Portanto, o debate entre Condorcet e Borda é um precursor de alguns debates modernos sobre como responder ao teorema de Arrow.

As origens desse debate precedem Condorcet e Borda. Na Idade Média, Ramon Llull (c. 1235-1315) propôs o método de agregação da votação majoritária em pares, enquanto Nicolas Cusanus (1401-1464) propôs uma variante da contagem de Borda (MCLEAN, 1990). Em 1672, o estadista e estudioso alemão Samuel von Pufendorf (1632-1694) comparou as regras de maioria simples, maioria qualificada e unanimidade e ofereceu uma análise da estrutura de preferências, que pode ser vista como precursora de descobertas posteriores, como é o caso do *pico único*, discutido abaixo. (GAERTNER, 2005).

No século XIX, o matemático e clérigo britânico Charles Dodgson (1832-1898), mais conhecido como Lewis Carroll, redescobriu de forma independente muitas das ideias já apresentadas por Condorcet e Borda e também desenvolveu uma teoria da representação proporcional. Foi em grande parte graças ao economista escocês Duncan Black (1908-1991) que as ideias teóricas da escolha social de Condorcet, Borda e Dodgson ganharam atenção da comunidade moderna de pesquisadores (MCLEAN; MCMILLAN; MONROE, 1995). Black também realizou várias descobertas relacionadas à votação majoritária, algumas das quais são discutidas abaixo.

Na França, George-Théodule Guilbaud (1966) escreveu um artigo importante, em 1952, mas muitas vezes negligenciado, que revisita a teoria da votação de Condorcet sob uma perspectiva lógica e provoca o desenvolvimento de uma literatura

francesa sobre o *efeito Condorcet*, o problema lógico subjacente ao paradoxo de Condorcet, que apenas recentemente recebeu maior atenção na teoria da escolha social anglófona (Monjardet, 2005). Para outras contribuições sobre a história da teoria da escolha social, consulte McLean, McMillan e Monroe (1996); McLean e Urken (1995); McLean e Hewitt (1994) e uma edição especial da revista *Social Choice and Welfare*, editada por Salles (2005).

2. Três argumentos formais para a regra da maioria

Para introduzir formalmente a teoria da escolha social, ajuda se considerar um problema simples de decisão: uma escolha coletiva entre duas alternativas.

2.1 O conceito de uma regra de agregação

Seja $N = \{1, 2, \dots, n\}$ um conjunto de indivíduos, onde $n \geq 2$. Os indivíduos devem escolher entre duas alternativas (candidatos, políticas públicas etc.). Cada indivíduo $i \in N$ deposita um voto, denotado v_i , onde:

- $v_i = 1$ representa um voto pela primeira alternativa,
- $v_i = -1$ representa um voto pela segunda alternativa, ou opcionalmente
- $v_i = 0$ representa uma abstenção (para simplificar, deixamos esta possibilidade de lado)

Uma combinação de votos dentre todos os indivíduos, $\langle v_1, v_2, \dots, v_n \rangle$, é chamada de *perfil*. Para qualquer perfil, o grupo busca chegar em uma decisão social v , onde:

- $v = 1$ representa uma decisão pela primeira alternativa,
- $v = -1$ representa uma decisão pela segunda alternativa, e
- $v = 0$ representa um empate.

Uma regra de agregação é uma função f que atribui a cada perfil $\langle v_1, v_2, \dots, v_n \rangle$ (em um domínio de perfis admissíveis) uma decisão social $v = f(v_1, v_2, \dots,$

v_n). Exemplos são:

Regra da maioria: Para cada perfil $\langle v_1, v_2, \dots, v_n \rangle$,

$$f(v_1, v_2, \dots, v_n) = \begin{cases} 1 & \text{se } v_1 + v_2 + \dots + v_n > 0 \\ & \text{("há mais 1s do que -1s");} \\ 0 & \text{se } v_1 + v_2 + \dots + v_n = 0 \\ & \text{("há tantos 1s quanto -1s");} \\ -1 & \text{se } v_1 + v_2 + \dots + v_n < 0 \\ & \text{("há mais -1s do que 1s").} \end{cases}$$

Ditadura: Para cada perfil $\langle v_1, v_2, \dots, v_n \rangle$,

$$f(v_1, v_2, \dots, v_n) = v_i,$$

onde $i \in N$ é um indivíduo antecedente fixo (o "ditador").

Regra da maioria ponderada: Para cada perfil $\langle v_1, v_2, \dots, v_n \rangle$,

$$f(v_1, v_2, \dots, v_n) = \begin{cases} 1 & \text{se } w_1v_1 + w_2v_2 + \dots + w_nv_n > 0, \\ 0 & \text{se } w_1v_1 + w_2v_2 + \dots + w_nv_n = 0, \\ -1 & \text{se } w_1v_1 + w_2v_2 + \dots + w_nv_n < 0, \end{cases}$$

onde $w_1, w_2, \dots, w_n$ são números reais, interpretados como "pesos dos votos" de n indivíduos.

Dois pontos sobre o conceito de regra de agregação são dignos de nota. Primeiro, sob a definição padrão, uma regra de agregação é definida *extensionalmente*, não *intensionalmente*: é um **mapeamento (relação funcional)** entre dados individuais iniciais e resultados coletivos, não um conjunto de **instruções explícitas** (uma **regra**, no sentido da linguagem comum) que poderia ser estendido a dados fora do domínio formal da função. Em segundo lugar, uma regra de agregação é definida para um conjunto fixo de indivíduos **N** e um problema de decisão fixo, de modo que a regra da maioria em um grupo de dois indivíduos é um objeto matemático diferente da regra da maioria em um grupo de três.

Para ilustrar, as Tabelas 1 e 2 mostram a regra da maioria para esses dois tamanhos de grupo como objetos extensionais. As linhas de cada tabela correspondem aos diferentes perfis possíveis de votos; a coluna final exhibe as decisões sociais

resultantes.

<i>Voto do indivíduo 1</i>	<i>Voto do indivíduo 2</i>	<i>Decisão coletiva</i>
1	1	1
1	-1	0
-1	1	0
-1	-1	-1

Tabela 1: Regra da maioria entre dois indivíduos

<i>Voto do indivíduo 1</i>	<i>Voto do indivíduo 2</i>	<i>Voto do indivíduo 3</i>	<i>Decisão coletiva</i>
1	1	1	1
1	1	-1	1
1	-1	1	1
1	-1	-1	-1
-1	1	1	1
-1	1	-1	-1
-1	-1	1	-1
-1	-1	-1	-1

Tabela 2: Regra da maioria entre três indivíduos

A forma atual de representar uma regra de agregação nos auxilia a perceber quantas regras de agregação possíveis existem (*vide* LIST, 2011). Suponha que existam k perfis no domínio dos dados admissíveis (no presente exemplo, $k = 2^n$, desde que cada um dos n indivíduos tenha duas escolhas, com abstenção não permitida). Suponha, ainda, que haja l decisões sociais possíveis para cada perfil (no exemplo, $l = 3$, permitindo empates). Então, existem l^k regras possíveis de agregação: a tabela relevante tem k linhas e, em cada linha, há l maneiras possíveis de especificar o registro final (a decisão coletiva). Portanto, o número de regras de agregação possíveis cresce exponencialmente com o número de perfis admissíveis e o número de resultados de decisão possíveis.

Para selecionar uma regra de agregação não arbitrariamente dentre a classe numerosa de regras possíveis, algumas restrições são necessárias. Agora considero três argumentos formais para a regra da maioria.

2.2 Um argumento procedimental para a regra da maioria

O primeiro argumento envolve a imposição de alguns requisitos “procedimentais” sobre a relação entre votos individuais e decisões sociais e mostra que a regra da maioria é a única regra de agregação que os satisfaz. May (1952) introduziu quatro desses requisitos:

Domínio universal (irrestrito): o domínio de dados iniciais admissíveis da regra de agregação consiste em todos os perfis logicamente possíveis de votos $\langle v_1, v_2, \dots, v_n \rangle$, onde cada $v_i \in \{-1, 1\}$.

Anonimato: para quaisquer perfis admissíveis $\langle v_1, v_2, \dots, v_n \rangle$ e $\langle w_1, w_2, \dots, w_n \rangle$ que são permutações um do outro (ou seja, um pode ser obtido do outro reordenando as entradas), a decisão social é a mesma, ou seja, $f(v_1, v_2, \dots, v_n) = f(w_1, w_2, \dots, w_n)$.

Neutralidade: para qualquer perfil admissível $\langle v_1, v_2, \dots, v_n \rangle$, se os votos para as duas alternativas forem invertidos, a decisão social também é invertida, ou seja, $f(-v_1, -v_2, \dots, -v_n) = -f(w_1, w_2, \dots, w_n)$.

Responsividade positiva: para qualquer perfil admissível $\langle v_1, v_2, \dots, v_n \rangle$, se alguns eleitores mudarem seus votos a favor de uma alternativa (digamos a primeira) e todos os outros votos permanecerem iguais, a decisão social não muda na direção contrária; se a decisão social foi um empate antes da mudança, o empate é desfeito na direção da mudança, ou seja, se $[w_i > v_i$ para alguns i e $w_j = v_j$ para todos os outros $j]$ e $f(v_1, v_2, \dots, v_n) = 0$ ou 1 , então $f(w_1, w_2, \dots, w_n) = 1$.

O pressuposto do domínio universal requer que a regra de agregação lide com qualquer nível de “pluralismo” em seus dados iniciais; o do anonimato exige que a regra trate todos os eleitores igualmente; e o da neutralidade exige que ela trate todas as alternativas igualmente; e a responsividade positiva requer que a decisão social seja uma função positiva da forma como as pessoas votam. May (1952) provou o seguinte:

Teorema: uma regra de agregação satisfaz os requisitos de domínio universal, anonimato, neutralidade e responsividade positiva se e somente se for a regra da maioria.

Além de fornecer um argumento para a regra da maioria com base em quatro desideratos procedimentais plausíveis, o teorema nos auxilia a caracterizar outras regras de agregação em termos de quais desideratos eles violam. Ditaduras e regras de maioria ponderada com pesos individuais desiguais violam o anonimato. Regras assimétricas de supermaioria (sob as quais uma supermaioria dos votos, como dois terços ou três quartos, é necessária para uma decisão a favor de uma das alternativas, enquanto a outra alternativa é a escolha automática padrão) violam a neutralidade. Isso pode às vezes ser justificável, por exemplo, quando há uma presunção a favor de uma alternativa, como a presunção de inocência em uma decisão de júri. As regras da supermaioria simétrica (sob as quais nenhuma alternativa é escolhida, a menos que seja apoiada por uma maioria suficientemente grande) violam a responsividade positiva. Um exemplo mais rebuscado de uma regra de agregação que viola responsividade positiva é a regra da maioria inversa (na qual a alternativa **rejeitada** pela maioria vence).

2.3 Um argumento epistêmico para a regra da maioria

O teorema do júri de Condorcet fornece um argumento consequentialista para a regra da maioria. O argumento é “epistêmico”, na medida em que a regra de agregação é interpretada como um dispositivo de rastreamento da verdade (GROFMAN; OWEN; FELD, 1983; LIST; GOODIN, 2001).

Suponha que o objetivo seja julgar algum fato independente de procedimento ou estado do mundo, denominado X . Em uma decisão de júri, o réu é ou culpado ($X = 1$) ou inocente ($X = -1$). Em uma decisão de um painel de especialistas sobre a segurança de alguma tecnologia, a tecnologia pode ser ou segura ($X = 1$) ou insegura ($X = -1$). O voto de cada indivíduo expressa um julgamento sobre aquele fato ou estado, e a decisão social representa o julgamento coletivo. O objetivo é chegar a um julgamento coletivo factualmente correto. Qual regra de agregação tem melhor desempenho em “rastrear a verdade” depende da relação entre os votos

individuais e o fato ou estado do mundo relevante.

Condorcet assumiu que cada indivíduo é mais capaz de fazer um julgamento correto (a suposição de competência) quando comparado a um julgamento aleatório e que os julgamentos de diferentes indivíduos são estocasticamente independentes, dado o estado do mundo (o pressuposto da independência). Formalmente, sejam $V_1, V_2, \dots, V_n$ (em letras maiúsculas) as variáveis aleatórias que geram os votos individuais específicos $v_1, v_2, \dots, v_n$ (em letras minúsculas), e seja $V = f(V_1, V_2, \dots, V_n)$ a variável aleatória resultante que representa a decisão social $v = f(v_1, v_2, \dots, v_n)$ sob uma dada regra de agregação f , tal como a regra da maioria. Os pressupostos de Condorcet podem ser apresentados da seguinte forma:

Competência: para cada indivíduo $i \in N$ e cada estado do mundo $x \in \{-1, 1\}$, $\Pr(V_i = x \mid X = x) = p > 1/2$, onde p é o mesmo entre indivíduos e estados.

Independência: os votos de diferentes indivíduos $V_1, V_2, \dots, V_n$ são independentes uns dos outros, condicionais a cada valor $x \in \{-1, 1\}$ de X .

Sob essas premissas, a votação majoritária é um bom rastreador da verdade:

Teorema (teorema do júri de Condorcet): para cada estado do mundo $x \in \{-1, 1\}$, a probabilidade de uma decisão majoritária correta, $\Pr(V = x \mid X = x)$, é maior do que a probabilidade de um voto individual correto, $\Pr(V_i = x \mid X = x)$, e converge para 1, conforme o número de indivíduos n aumenta.²⁵

A primeira conjunção (“é maior do que a probabilidade de um voto individual”) é a **conclusão não assintótica**, a segunda (“converge para 1”) é a **conclusão assintótica**. É possível ainda mostrar que, se os dois estados do mundo têm uma probabilidade anterior igual (i.e., $\Pr(X = 1) = \Pr(X = -1) = 1/2$), a regra da maioria

²⁵ Quando n é par, a primeira parte do teorema só é válida para tamanhos de grupos n acima de um certo limite inferior (que depende de p), devido à possibilidade de empates majoritários. Quando n é ímpar, vale para qualquer $n > 1$.

é a mais confiável de todas as regras de agregação, maximizando $\Pr(V = X)$ (como em BEN-YASHAR; NITZAN, 1997).

Embora o teorema do júri seja frequentemente invocado para estabelecer os méritos epistêmicos da democracia, seus pressupostos são altamente idealistas. O pressuposto de competência não é uma afirmação conceitual, mas empírica e depende do problema de decisão dado. Embora uma competência individual média (não necessariamente igual) acima de $1/2$ possa ser suficiente para a conclusão de Condorcet (GROFMAN; OWEN; FELD, 1983; BOLAND, 1989; KANAZAWA, 1998),²⁶ o teorema deixa de ser válido se as decisões dos indivíduos forem aleatórias (nem melhor nem pior do que cara ou coroa) ou se forem piores do que aleatórias ($p < 1/2$). No último caso, a probabilidade de uma decisão majoritária correta é *menor* do que a probabilidade de cada indivíduo de apresentar um voto correto e converge para 0, conforme o tamanho do júri aumenta. A conclusão do teorema também pode ser enfraquecida em casos menos extremos (BEREND; PAROUSH, 1998), por exemplo quando a confiabilidade de cada indivíduo, embora acima de $1/2$, é uma função exponencialmente decrescente que se aproxima de $1/2$ com o aumento do tamanho do júri (LIST, 2003a).

De forma similar, a veracidade da suposição de independência depende do problema de decisão em questão. Embora a conclusão de Condorcet seja robusta à presença de algumas interdependências entre votos individuais, a estrutura dessas interdependências é importante (BOLAND, 1989; LADHA, 1992; ESTLUND, 1994; DIETRICH; LIST, 2004; BEREND; SAPIR, 2007; DIETRICH; SPIEKERMANN, 2013). Se todos os votos dos indivíduos estão perfeitamente correlacionados uns com os outros ou imitam um pequeno número de líderes de opinião, o julgamento coletivo não é mais confiável do que o julgamento entre um pequeno número de indivíduos independentes.

Redes bayesianas, conforme empregadas no trabalho de Pearl sobre causalidade (2000), têm sido empregadas para modelar os efeitos das dependências dos votantes no teorema do júri e para distinguir entre variantes mais fortes e mais

²⁶ Se diferentes indivíduos têm diferentes níveis *conhecidos* de confiabilidade, a votação por maioria ponderada supera a votação por maioria simples na maximização da probabilidade de uma decisão correta, com o peso do voto de cada indivíduo proporcional a $\log(p / (1 - p))$, onde p é a confiabilidade do indivíduo conforme definida acima (SHAPLEY; GROFMAN, 1984; GROFMAN; OWEN; FELD, 1983; BEN-YASHAR; NITZAN, 1997).

fracas de independência condicional (DIETRICH; LIST, 2004; DIETRICH; SPIEKERMANN, 2013). Dietrich (2008) argumentou que os dois pressupostos de Condorcet nunca são simultaneamente justificados no sentido de que, mesmo quando ambos são verdadeiros, não se pode obter evidências em favor de ambos ao mesmo tempo.

Finalmente, a pesquisa em teoria dos jogos desafia uma suposição implícita do teorema do júri, qual seja a que os eleitores sempre revelarão seus julgamentos com veracidade. Mesmo se todos os eleitores preferirem um julgamento coletivo correto ao incorreto, eles ainda podem ter incentivos para representar mal seus julgamentos individuais. Isso pode acontecer quando, condicionado ao evento de ser fundamental para o resultado, um votante tem maior expectativa de alcançar um julgamento coletivo correto votando **contra** seu próprio julgamento privado do que em consonância com ele (AUSTIN-SMITH E BANKS, 1996; FEDDERSEN; PESENDORFER, 1998).

2.4 Um argumento utilitarista para a regra da maioria

Outro argumento consequencialista para o governo da maioria é utilitarista ao invés de epistêmico. Ele não requer a existência de um fato ou estado do mundo independentes que a decisão coletiva supostamente rastreia. Suponha que cada votante obtenha alguma utilidade da decisão coletiva, o que depende de a decisão corresponder ao seu voto (preferência): especificamente, cada votante obtém uma utilidade de 1 de uma correspondência entre seu voto e o resultado coletivo e uma utilidade de 0 de uma incompatibilidade²⁷. O teorema de Rae-Taylor afirma, assim, que se cada indivíduo tem uma probabilidade anterior igual de preferir cada uma das duas alternativas, a regra da maioria maximiza a utilidade esperada de cada indivíduo (*vide*, por exemplo, MUELLER, 2003).

De forma relacionada, a regra da maioria minimiza o número de votantes frustrados (definidos como votantes do lado perdedor) e maximiza a utilidade total entre os votantes. Brighthouse e Fleurbaey (2010) generalizam esse resultado. Definamos o *interesse* do votante i na decisão, d_i , como a diferença de utilidade

²⁷ Opcionalmente, pode-se estipular que a utilidade de um empate é 1/2.

entre seu resultado preferido e seu resultado diferido. O teorema de Rae-Taylor baseia-se em uma suposição implícita de **interesses** iguais, ou seja, $d_i = 1$ para cada $i \in N$. Brighthouse e Fleurbaey mostram que quando os interesses podem variar entre os votantes, a utilidade total é maximizada não pela regra da maioria, mas por uma regra de maioria ponderada, em que o peso w_i do voto de cada indivíduo i é proporcional a seu interesse d_i .

3. Agregação de preferências

No cerne da teoria da escolha social está a análise da agregação de preferências, compreendida como a agregação de ordenações de preferência de duas ou mais alternativas sociais de diversos indivíduos em uma única ordenação de preferência (ou escolha) coletiva sobre essas alternativas.

O modelo básico é o seguinte. Novamente, considere um conjunto $N = \{1, 2, \dots, n\}$ de indivíduos ($n \geq 2$). Seja $X = \{x, y, z, \dots\}$ um conjunto de alternativas sociais, por exemplo, mundos possíveis, plataformas de políticas públicas, candidatos eleitorais ou alocações de bens. Cada indivíduo $i \in N$ tem uma **ordem de preferência** R_i sobre essas alternativas: uma relação binária completa e transitiva de X .²⁸ Para qualquer $x, y \in X$, $xR_i y$ significa que o indivíduo i prefere fracamente x a y . Escrevemos $xP_i y$ se $xR_i y$ e não $yR_i x$ ("indivíduo i prefere estritamente x a y "), e $xI_i y$ se $xR_i y$ e $yR_i x$ ("indivíduo i é indiferente entre x e y ").

Uma combinação de ordens de preferência entre os indivíduos, $\langle R_1, R_2, \dots, R_n \rangle$, é chamada de *perfil*. Uma *regra de agregação de preferência*, F , é uma função que atribui a cada perfil $\langle R_1, R_2, \dots, R_n \rangle$ (em algum domínio de perfis admissíveis) uma relação de preferência social $R = F(R_1, R_2, \dots, R_n)$ em X . Quando F é claro no contexto, simplesmente escrevemos R para a relação de preferência social correspondente a $\langle R_1, R_2, \dots, R_n \rangle$.

Para qualquer $x, y \in X$, xRy significa que x é socialmente fracamente preferível a y . Também escrevemos xPy se xRy e não yRx (" x é estritamente socialmente preferível a y "), e xIy se xRy e yRx (" x e y estão socialmente empatados"). Por generalidade, o requisito de que R seja completo e transitivo não está embutido

²⁸ Completude requer que, para qualquer $x, y \in X$, $xRiy$ ou $yRix$, e transitividade requer que, para qualquer $x, y, z \in X$, se $xRiy$ e $yRiz$, então $xRiz$.

na definição de uma regra de agregação de preferências.

O exemplo paradigmático de uma regra de agregação de preferências é a votação majoritária em pares, conforme discutido por Condorcet. Aqui, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se pelo menos tantos indivíduos têm xRy quanto têm yR_x , formalmente $|\{i \in N: xR_i y\}| \geq |\{i \in N: yR_i x\}|$. Conforme visto, isso não garante preferências sociais transitivas.²⁹

Quão frequentes são as preferências intransitivas da maioria? É possível mostrar que a proporção de perfis de preferência (entre todos os possíveis) que levam a preferências majoritárias cíclicas aumenta com o número de indivíduos (n) e o número de alternativas ($|X|$). Se todos os perfis de preferência possíveis são igualmente prováveis de ocorrer (o chamado cenário de “cultura imparcial”), então os ciclos majoritários devem ser prováveis em grandes eleitorados (GEHRLEIN, 1983). (O trabalho técnico ainda distingue entre “ciclos superiores” e ciclos abaixo de uma possível alternativa vencedora de Condorcet). No entanto, a probabilidade de ciclos pode ser significativamente menor sob certos desvios sistemáticos, mesmo pequenos, de uma cultura imparcial (LIST; GOODIN, 2001, apêndice 3; TSETLIN, REGENWETTER; GROFMAN, 2003; REGENWETTER *et al.*, 2006).

3.1 Teorema de Arrow

Abstraindo da votação majoritária em pares, Arrow introduziu as seguintes condições em uma regra de agregação de preferência, F .

Domínio universal: o domínio de F é o conjunto de todos os perfis logicamente possíveis de ordenações de preferência individuais completas e transitivas.

Ordenação: para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ no domínio de F , a relação de preferência social R é completa e transitiva.

²⁹ No exemplo clássico, existem três indivíduos com ordenações de preferência xP_1yP_1z , yP_2zP_2x e zP_3xP_3y relativamente a três alternativas x , y e z . As preferências majoritárias resultantes são cíclicas: temos xPy , yPz além de zPx .

Princípio fraco de Pareto: para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ no domínio de F , se para todo $i \in N$ $xP_i y$, então xPy .

Independência das alternativas irrelevantes: para quaisquer dois perfis $\langle R_1, R_2, \dots, R_n \rangle$ e $\langle R_1^*, R_2^*, \dots, R_n^* \rangle$ no domínio de F e qualquer $x, y \in X$, se para todos $i \in N$ a ordenação de R_i entre x e y coincide com a ordenação de R_i^* entre x e y , então xRy se e somente se xR^*y .

Não-ditadura: não existe um indivíduo $i \in N$ tal que, para todos $\langle R_1, R_2, \dots, R_n \rangle$ no domínio de F e todo $x, y \in X$, $xP_i y$ implica xPy .

O domínio universal requer que a regra de agregação lide com qualquer nível de “pluralismo” em seus dados iniciais. A ordenação exige que ela produza preferências sociais “racionais”, evitando os ciclos de Condorcet. O princípio fraco de Pareto requer que, quando todos os indivíduos preferirem estritamente a alternativa x à alternativa y , também o faça a sociedade. A independência de alternativas irrelevantes requer que a preferência social entre quaisquer duas alternativas x e y dependa apenas das preferências individuais entre x e y , não das preferências dos indivíduos sobre outras alternativas. Não-ditadura exige que não haja “ditador”, alguém que sempre determina a preferência social, independentemente das preferências de outros indivíduos. (Note que a votação majoritária em pares satisfaz todas essas condições, exceto a ordenação).

Teorema (ARROW, 1951, 1963): se $|X| > 2$, não existe regra de agregação de preferência que satisfaça domínio universal, ordenação, princípio fraco de Pareto, independência de alternativas irrelevantes e não-ditadura.

É evidente que esse resultado é transportado para a agregação de outros tipos de ordenações, distintas das ordenações de preferência, como (i) ordenações de crença sobre várias hipóteses (crenças ordinais), (ii) diversos critérios que um

único tomador de decisão pode usar para, tudo o mais considerado, gerar uma ordenação de várias opções de decisão e (iii) ordenações de valor conflitantes a serem reconciliadas.

Exemplos de outros problemas de agregação aos quais o teorema de Arrow tem sido aplicado incluem: problemas de agregação intrapessoal (por exemplo/ e.g. MAY, 1954; HURLEY, 1985), agregação de restrições na teoria de otimalidade na área de linguística (por exemplo/ e.g. HARBOR; LIST, 2000), teoria da escolha (por exemplo/ e.g. OKASHA, 2011; MORREAU, no prelo), amalgamação de evidências (por exemplo/ e.g. STEGENGA, 2013) e a agregação de múltiplas ordens de similaridade em uma ordem de similaridade “com tudo considerado” (por exemplo/ e.g. MORREAU, 2010; KROEDEL; HUBER, 2013). Em cada caso, a plausibilidade do teorema de Arrow depende da plausibilidade específica ao caso do arcabouço ordinalista de Arrow e das condições do teorema.

Geralmente, se considerarmos o arcabouço de Arrow apropriado e suas condições indispensáveis, o teorema de Arrow cria um sério desafio. Para evitá-lo, devemos relaxar pelo menos uma das cinco condições ou desistir da restrição dos dados iniciais da regra de agregação para ordenações e defender o uso de dados iniciais mais ricos, conforme discutido na Seção 4.

3.2 Regras de agregação de preferências não ditatoriais

3.2.1 Relaxando o domínio universal

Uma maneira de evitar o teorema de Arrow é relaxar o domínio universal. Se for requerido à regra de agregação aceitar como dados iniciais apenas perfis de preferência que satisfaçam certas condições de “coesão”, então regras de agregação como votação majoritária em pares produzirão preferências sociais completas e transitivas. A condição de coesão mais conhecida é chamada *pico único* (BLACK, 1948).

Um perfil $\langle R_1, R_2, \dots, R_n \rangle$ possui *pico único* se as alternativas puderem ser alinhadas da “esquerda” para a “direita” (por exemplo, em alguma dimensão cognitiva ou ideológica), tal que cada indivíduo tenha uma posição de maior preferência nesse alinhamento com preferência decrescente conforme as alternativas ficam mais distantes (em qualquer direção) da posição de maior preferência. Formalmente,

isso requer a existência de uma ordem linear Ω em X de modo que, para cada trio de alternativas $x, y, z \in X$, se y estiver entre x e z em Ω , não é o caso de $xR_y y$ e $zR_x x$ (isso exclui uma cava entre x e z , em y). O *pico único* é plausível em alguns contextos democráticos. Se as alternativas em X são diferentes taxas de tributos, por exemplo, cada indivíduo pode ter uma taxa de maior preferência (que será menor para um indivíduo libertário do que para um socialista) e ter menor preferência por outras taxas à medida que se distanciam mais do ideal.

Black (1948) provou que, se o domínio da regra de agregação é restrito ao conjunto de todos os perfis de ordenações de preferência individuais que satisfazem a condição de *pico único*, os ciclos de maioria não podem ocorrer, e a alternativa de maior preferência do indivíduo mediano em relação ao alinhamento esquerda-direita é uma vencedora de Condorcet (assumindo que n é ímpar). Assim, a votação majoritária em pares satisfaz as demais condições de Arrow.

Outras condições de restrição de domínio com implicações semelhantes incluem a de *cava-única*, uma imagem geométrica de espelho do *pico único* (INADA, 1964), *separabilidade em dois grupos* (*ibid.*) e a *ausência do quadrado latino* (WARD, 1965), sendo as duas últimas condições combinatórias mais complexas (para uma revisão, ver Gaertner, 2001). Sen (1966) mostrou que todas essas condições implicam uma condição mais fraca, *restrição de valor por trios*. Essa condição requer que, para cada trio de alternativas $x, y, z \in X$, exista uma alternativa em $\{x, y, z\}$ e uma ordenação $r \in \{1, 2, 3\}$ tal que nenhum indivíduo a ordene no r -ésimo lugar entre x, y e z . Por exemplo, todos os indivíduos podem concordar que y não é a última ($r = 3$) entre x, y e z . A *restrição de valor por trios* é suficiente para preferências transitivas majoritárias (para uma demonstração simples do teorema de Sen, consulte Elsholtz e List, 2005).

Tem havido muita discussão sobre se, e sob que condições, as preferências do mundo real caem em tal domínio restrito. Foi sugerido, por exemplo, que a deliberação em grupo pode induzir preferências de *pico único*, ao direcionar os participantes a se concentrarem em uma dimensão cognitiva ou ideológica compartilhada (MILLER, 1992; KNIGHT; JOHNSON, 1994; DRYZEK; LIST, 2003). Evidência experimental de pesquisas de opinião deliberativas é consistente com essa hipótese (LIST, LUSKIN, FISHKIN, MCLEAN, 2013), embora sejam necessárias mais pesquisas empíricas.

3.2.2 Relaxando a ordem

Espera-se normalmente que regras de agregação de preferência produzam ordenações como resultados de suas aplicações, mas às vezes pode-se exigir apenas ordenações parciais ou relações binárias não totalmente transitivas. Uma regra de agregação que produz preferências sociais transitivas mas frequentemente incompletas, é o *procedimento de dominância de Pareto*: aqui, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se, para todo $i \in N$, $xP_i y$. Uma regra de agregação que produz preferências sociais completas mas frequentemente intransitivas é o *procedimento de extensão de Pareto*: aqui, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se não for o caso que, para todo $i \in N$, $yP_i x$. Ambas as regras têm espírito de unanimidade, dando a cada indivíduo poder de veto contra a *presença* de uma preferência social fraca por x a y ou contra sua ausência.

Gibbard (1969) provou que, mesmo se substituirmos o requisito de transitividade pelo que ele chamou de *quase transitividade*, as possibilidades de agregação resultantes ainda são muito limitadas. Chamemos uma relação de preferência R de *quase transitiva* se a relação estrita induzida P for transitiva (enquanto a relação de indiferença I não precisa ser transitiva). Chamemos uma regra de agregação de *oligárquica* se houver um subconjunto $M \subseteq N$ (os “oligarcos”) tal que (i) se, para todo $i \in M$, $xP_i y$, então xPy , e (ii) se, para algum $i \in M$, $xP_i y$ e, então, xRy . O procedimento de extensão de Pareto é um exemplo de regra de agregação oligárquica com $M = N$. Em uma oligarquia, os oligarcas são decisivos em conjunto e têm poder individual de veto. Gibbard provou o seguinte:

Teorema (GIBBARD, 1969): se $|X| > 2$, não há regra de agregação de preferência que satisfaça domínio universal, quase transitividade e completude das preferências sociais, o princípio fraco de Pareto, independência de alternativas irrelevantes e não-oligarquia.

3.2.3 Relaxando o princípio fraco de Pareto

O princípio fraco de Pareto é difícil de abandonar. Um caso em que podemos abdicar dele é o da unanimidade espúria, em que uma preferência unânime por x

relativamente a y é baseada em razões mutuamente inconsistentes (MONGIN, 1997; GILBOA, SAMET; SCHMEIDLER, 2004). Dois homens podem preferir lutar um duelo (alternativa x) a não lutar (alternativa y), porque ambos superestimam suas chances de vitória. Pode não haver atribuição de probabilidade mutuamente aceitável sobre os resultados possíveis do duelo (ou seja, quem venceria) que poderia “racionalizar” a preferência unânime de x em detrimento de y . Nesse caso, a preferência unânime é um indicador ruim de preferência social. Este exemplo, entretanto, depende do fato de que as alternativas de lutar e não lutar não são resultados totalmente especificados, mas perspectivas incertas. Pode-se afirmar que o princípio fraco de Pareto é mais plausível em casos sem incerteza.

Uma regra de agregação que se torna possível quando o princípio fraco de Pareto é abandonado é um *procedimento imposto*, em que, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$, a relação de preferência social R é uma ordenação $R_{imposta}$ previamente fixada (“imposta”) das alternativas. Embora completamente não-responsiva às preferências individuais, essa regra de agregação satisfaz o restante das condições de Arrow.

Sen (1970a) ofereceu outra crítica ao princípio fraco de Pareto, mostrando que ele entra em conflito com um princípio “liberal”. Aqui, interpretamos a regra de agregação como um método que um planejador social pode usar para ordenar as alternativas sociais em uma ordem de bem-estar social. Suponha que cada indivíduo na sociedade detenha alguns direitos básicos, no sentido de que sua preferência às vezes é socialmente decisiva (ou seja, não pode ser anulada pelas preferências dos outros). Tanto Lewd quanto Prude, por exemplo, devem individualmente decidir sobre se lerão um livro específico, *O Amante de Lady Chatterley*³⁰.

Liberalismo mínimo: existem ao menos dois indivíduos distintos $i, j \in N$ que são cada um decisivos em pelo menos um par de alternativas; ou seja, há pelo menos um par de alternativas $x, y \in X$ tal que, para cada perfil $\langle R_1, R_2, \dots, R_n \rangle$, $xP_i y$ implica xPy e $yP_j x$ implica yPx , e ao menos um par de alternativas $x^*, y^* \in X$

³⁰ N.T.: Sen baseia-se aqui em um livro que causou, em meados do século passado, polêmicas por conter cenas eróticas. Os nomes dos personagens, no exemplo, “Prude” e “Lewd”, remetem a uma possível disputa nesse contexto, significando, em inglês, “Pudico” e “Lascivo”.

tal que, para cada perfil $\langle R_1, R_2, \dots, R_n \rangle$, $x^*P_j y^*$ implica $x^*P y^*$ e $y^*P x^*$ implica $y^*P x^*$.

Sen pede que imaginemos que Lewd possui maior preferência que Prude leia o livro (alternativa x), prefere em segundo lugar ler o livro sozinho (alternativa y) e o que ele menos prefere é que nenhum dos dois leia o livro (z). Prude tem como maior preferência que nenhum dos dois leia o livro (z), prefere em segundo lugar que ele mesmo leia o livro (x) e tem como menor preferência que Lewd leia o livro (y). Assumindo que Lewd é decisivo sobre o par y e z , a sociedade deve preferir y a z . Supondo que Prude é decisivo sobre o par x e z , a sociedade deve preferir z a x . Mas, dado que tanto Lewd quanto Prude preferem x a y , o princípio fraco de Pareto (aplicado a $N = \{\text{Lewd}, \text{Prude}\}$) implica que a sociedade deve preferir x a y . Portanto, estamos diante de um ciclo de preferência social. Sen chamou esse problema de “paradoxo liberal” e generalizou-o da seguinte maneira.

Teorema (SEN, 1970a): Não existe regra de agregação de preferências que satisfaça domínio universal, aciclicidade das preferências sociais, princípio fraco de Pareto e liberalismo mínimo.

O resultado sugere que, se desejarmos respeitar os direitos individuais, às vezes teremos que sacrificar a eficiência paretiana. Uma conclusão alternativa é que o princípio de Pareto fraco pode ser tornado compatível com o liberalismo mínimo apenas quando o domínio dos perfis de preferência admissíveis é adequadamente restrito, por exemplo, às preferências que são “tolerantes” ou não “intrusivas” (BLAU, 1975; CRAVEN, 1982; GIGLIOTTI, 1986; SEN, 1983). As preferências de Lewd e Prude no exemplo de Sen são “intrusivas”.

Vários autores contestaram a relevância do resultado de Sen, criticando, no entanto, a formalização de direitos por ele realizada (*vide* GAERTNER, PATTANAİK; SUZUMURA, 1992; DOWDING; VAN HEES, 2003).

3.2.4 Relaxando a independência de alternativas irrelevantes

Uma maneira comum de obter regras possíveis de agregação de preferências é renunciando à independência de alternativas irrelevantes. Quase todos os métodos

conhecidos de votação com três ou mais alternativas que envolvem alguma forma de votação preferencial (com os eleitores sendo solicitados a expressar ordens de preferência total ou parcial) violam essa condição.

Um exemplo padrão é a *regra de pluralidade*: aqui, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $|\{i \in N: \text{para todo } z \neq x, xPz\}| \geq |\{i \in N: \text{para todo } z \neq y, yPz\}|$. Informalmente, as alternativas são socialmente ordenadas de acordo com quantos indivíduos possuem maior preferência por cada uma delas. A regra da pluralidade evita o paradoxo de Condorcet, mas esbarra em outros problemas. Sobretudo, uma alternativa que a maioria não prefere a todas as outras alternativas pode vencer sob a regra da pluralidade: se 34% dos eleitores ordenam x acima de y acima de z , 33% ordenam y acima de z acima de x , e 33% ordenam z acima de y acima de x , a regra da pluralidade ordena x acima de y e de z , enquanto a votação majoritária em pares ordenaria y acima de z acima de x (y é o vencedor de Condorcet). Ao desconsiderar as alternativas de ordenação inferior dos indivíduos, a regra da pluralidade também viola o princípio fraco de Pareto. No entanto, a regra da pluralidade pode ser plausível em “ambientes informativos restritos”, em que o procedimento de votação coleta informações apenas sobre as preferências maiores dos eleitores, não sobre suas ordenações completas de preferência. Aqui, a regra da pluralidade satisfaz variantes generalizadas das quatro condições de May apresentadas acima (GOODIN; LIST, 2006).

Um segundo exemplo de uma regra de agregação de preferência que viola a independência de alternativas irrelevantes é a *contagem de Borda* (vide SAARI, 1990). Aqui, para qualquer perfil $\langle R_1, R_2, \dots, R_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $\sum_{i \in N} |\{z \in X : xRz\}| \geq \sum_{i \in N} |\{z \in X : yRz\}|$. Informalmente, cada votante atribui uma pontuação a cada alternativa dependendo do lugar que ela ocupa em sua ordenação de preferências. A alternativa de maior preferência obtém uma pontuação k (onde $k = |X|$), a segunda alternativa de maior preferência uma pontuação de $k - 1$, a terceira alternativa de maior preferência uma pontuação de $k - 2$ e assim por diante. As alternativas são então socialmente ordenadas em termos das somas de suas pontuações entre os votantes: a alternativa com a maior soma total é a primeira, a alternativa com a segunda maior soma total a seguir, e assim por diante.

Para perceber como isso viola a independência de alternativas irrelevantes, considere os dois perfis de ordenações de preferência individuais sobre quatro alternativas (x, y, z, w) nas Tabelas 3 e 4.

	Indivíduo 1	Indivíduos 2 a 7	Indivíduos 8 a 15
1ª preferência	y	x	z
2ª preferência	x	z	x
3ª preferência	z	w	y
4ª preferência	w	y	w

Tabela 3: um perfil de ordens de preferência individuais.

	Indivíduo 1	Indivíduos 2 a 7	Indivíduos 8 a 15
1ª preferência	x	x	z
2ª preferência	y	z	x
3ª preferência	w	w	y
4ª preferência	z	y	w

Tabela 4: Um perfil ligeiramente modificado de ordens de preferência individuais

Na Tabela 3, as pontuações de Borda das quatro alternativas são:

- $x: 9 \cdot 3 + 6 \cdot 4 = 51,$
- $y: 1 \cdot 4 + 6 \cdot 1 + 8 \cdot 2 = 26,$
- $z: 1 \cdot 2 + 6 \cdot 3 + 8 \cdot 4 = 52,$
- $w: 1 \cdot 1 + 6 \cdot 2 + 8 \cdot 1 = 21,$

resultando em uma preferência social por z a x, x a y, y a w. Na Tabela 4, as pontuações de Borda são:

- $x: 7 \cdot 4 + 8 \cdot 3 = 52,$
- $y: 1 \cdot 3 + 6 \cdot 1 + 8 \cdot 2 = 25,$
- $z: 1 \cdot 1 + 6 \cdot 3 + 8 \cdot 4 = 51,$
- $w: 7 \cdot 2 + 8 \cdot 1 = 22,$

resultando em uma preferência social por x a z a y a w. A única diferença entre os dois perfis está na ordem de preferência do Indivíduo 1 e, mesmo aqui, não há mudança na ordenação relativa de x e z. Apesar das preferências individuais idênticas entre x e z nas Tabelas 3 e 4, a preferência social entre x e z é invertida, uma violação da independência de alternativas irrelevantes.

Essas violações são comuns nas regras de votação do mundo real e tornam a agregação de preferências potencialmente vulnerável à votação estratégica e /

ou definição de agenda estratégica. Ilustro isso no caso da votação estratégica.

3.3 Teorema de Gibbard-Satterthwaite

Até o momento, discutimos *regras de agregação de preferências*, que mapeiam perfis de ordenações de preferência individuais para relações de preferência social. Agora consideraremos as *regras de escolha social*, cujo resultado, diferentemente, é uma ou várias alternativas vencedoras. Formalmente, uma *regra de escolha social*, f , é uma função que atribui a cada perfil $\langle R_1, R_2, \dots, R_n \rangle$ (em algum domínio de perfis admissíveis) um conjunto de escolha social $f(R_1, R_2, \dots, R_n) \subseteq X$. Uma regra de escolha social f pode ser derivada de uma regra de agregação de preferência F , ao definir $f(R_1, R_2, \dots, R_n) = \{x \in X : \text{para todo } y \in X, xRy\}$ onde $R = F(R_1, R_2, \dots, R_n)$; o inverso geralmente não é válido. Chamamos o conjunto de alternativas às vezes escolhidas de intervalo de f .³¹

O *critério vencedor de Condorcet* define uma regra de escolha social, onde, para cada perfil $\langle R_1, R_2, \dots, R_n \rangle$, $f(R_1, R_2, \dots, R_n)$ contém todas as alternativas em X que ganha ou pelo menos empata com todas as outras na votação majoritária em pares. Como mostra o paradoxo de Condorcet, isso pode produzir um conjunto de escolha vazio. Em contraste, a regra da pluralidade e a contagem de Borda induzem regras de escolha social que sempre produzem conjuntos de escolha não-vazios. Eles também satisfazem as seguintes condições básicas (a última para $|X| \geq 3$):

Domínio universal: o domínio de f é o conjunto de todos os perfis logicamente possíveis de ordenações de preferência individuais completas e transitivas.

Não-ditadura: não existe um indivíduo $i \in N$ tal que, para todo $\langle R_1, R_2, \dots, R_n \rangle$ no domínio de f e todo x

³¹ Formalmente, $\{x \in X : x \in f(R_1, R_2, \dots, R_n) \text{ para algum } \langle R_1, R_2, \dots, R_n \rangle \text{ no domínio de } f\}$.

no intervalo de f , yR_ix onde $y \in f(R_1, R_2, \dots, R_n)$.³²

A restrição de intervalo: o intervalo de f contém pelo menos três alternativas distintas (e idealmente todas as alternativas em X).

Quando complementadas com um critério de desempate apropriado, as regras da pluralidade e de Borda podem ainda ser tornadas “determinadas”:

Determinação: a regra de escolha social f sempre produz uma alternativa vencedora única – um conjunto de escolha de um único elemento (então, escrevemos $x = f(R_1, R_2, \dots, R_n)$ para denotar a alternativa vencedora para o perfil $\langle R_1, R_2, \dots, R_n \rangle$).

Surpreendentemente, esta lista de condições está em conflito com o seguinte requisito adicional.

Resistência à estratégia: não existe um perfil $\langle R_1, R_2, \dots, R_n \rangle$ no domínio de f onde f é manipulável por algum indivíduo $i \in N$, onde a *manipulabilidade* significa o seguinte: se i apresenta uma ordem falsa de preferência $R'_i (\neq R_i)$, o vencedor é uma alternativa y' que i estritamente prefere (de acordo com R_i) à alternativa y que ganharia se i submetesse a ordem de preferência verdadeira R_i .³³

Teorema (GIBBARD, 1973; SATTERTHWAIT, 1975): não existe regra de escolha social que satisfaça domínio universal, não-ditadura, restrição de intervalo, determinação e resistência à estratégia.

³² Para os presentes propósitos, pode-se estipular que a última cláusula (para todo x no intervalo de f , yR_ix em que $y \in f(R_1, R_2, \dots, R_n)$) é violada se $f(R_1, R_2, \dots, R_n)$ é vazio.

³³ Formalmente, $y'P_iy$, em que $y' = f(R_1, \dots, R'_i, \dots, R_n)$ e $y = f(R_1, \dots, R_i, \dots, R_n)$, assumindo que $\langle R_1, \dots, R'_i, \dots, R_n \rangle$ está no domínio de f . A definição pressupõe que os conjuntos de escolha social para os perfis $\langle R_1, \dots, R_i, \dots, R_n \rangle$ e $\langle R_1, \dots, R'_i, \dots, R_n \rangle$ sejam únicos.

Esse resultado gera questões importantes sobre as compensações (*trade-offs*) entre diferentes requisitos em uma regra de escolha social. Uma ditadura, que sempre escolhe a alternativa preferida do ditador, é trivialmente resistente à estratégia. O ditador obviamente não tem incentivos para votar estrategicamente, e ninguém mais o faz, já que o resultado depende apenas do ditador.

Para ver que a contagem de Borda não é resistente à estratégia, lembre-se do exemplo das Tabelas 3 e 4 acima. Se o Indivíduo 1 na Tabela 3 realmente enviar a ordenação de preferências $yP_1xP_1zP_1w$, o vencedor de Borda será z , como vimos. Se o Indivíduo 1 apresentar falsamente a ordenação de preferências $xP_1yP_1wP_1z$, como na Tabela 4, o vencedor de Borda é x . Porém o Indivíduo 1 prefere x a z de acordo com sua ordem verdadeira de preferências (na Tabela 3) e, portanto, ele ou ela tem um incentivo para votar estrategicamente.

Moulin (1980) mostrou que quando o domínio da regra de escolha social é restrito a perfis de preferência de pico único, a votação majoritária em pares e outros esquemas de “votação mediana” podem satisfazer o resto das condições do teorema de Gibbard-Satterthwaite. De forma similar, quando as decisões coletivas são restritas a escolhas binárias apenas, o que equivale a eliminar a condição de restrição de intervalo, a votação majoritária satisfaz o resto das condições. Outras possíveis rotas de fuga do teorema se abrem se a condição de determinação for abandonada. No caso limite em que todas as alternativas são sempre escolhidas, as outras condições são satisfeitas de forma vazia.

O requisito de ser resistente à estratégia também foi desafiado. Uma linha de argumento é que, mesmo quando existem incentivos estratégicos no sentido técnico do teorema de Gibbard-Satterthwaite, os indivíduos não necessariamente agirão sobre eles. Eles exigiriam informações detalhadas sobre as preferências dos outros e poder computacional suficiente para descobrir quais seriam as preferências ótimas estrategicamente modificadas. Nenhuma das demandas é geralmente atendida. Bartholdi, Tovey e Trick (1989) mostraram que, devido à complexidade computacional, algumas regras de escolha social são resistentes à manipulação estratégica: pode ser um problema NP-difícil para um eleitor ou eleitora determinar como votar estrategicamente. Nesse sentido, Harrison e McDaniel (2008) fornecem evidências experimentais sugerindo que a “regra de Kemeny”, uma extensão da votação majoritária em pares projetada para evitar ciclos de Condorcet, é “compatível com incentivos comportamentais”: isto é, a manipulação estratégica é computacionalmente difícil.

Dowding e van Hees (2008) argumentaram que nem todas as formas de votação estratégica são normativamente problemáticas. Eles distinguem entre formas “sinceras” e “não sinceras” de manipulação e argumentam que apenas a última forma, mas não a primeira, é normativamente problemática. A manipulação sincera ocorre quando um eleitor ou eleitora (i) vota em uma alternativa de compromisso cujas chances de vitória aumentam e (ii) genuinamente prefere essa alternativa de compromisso à alternativa que, de outra forma, venceria. Por exemplo, na eleição presidencial dos EUA de 2000, os apoiadores de Ralph Nader (um candidato de um terceiro partido com poucas chances de vitória) que votaram em Al Gore para aumentar suas chances de derrotar George W. Bush se envolveram em manipulação sincera conforme (i) e (ii). A regra da pluralidade é suscetível à manipulação sincera, mas invulnerável à manipulação insincera.

4. Agregação de bem-estar

Uma suposição implícita até agora tem sido que as preferências são ordinais e não interpessoalmente comparáveis: as ordenações de preferências não contêm informações sobre a força da preferência de cada indivíduo ou sobre como comparar as preferências de diferentes indivíduos entre si. Afirmarções como “Indivíduo 1 prefere a alternativa x mais do que o Indivíduo 2 prefere a alternativa y ” ou “Indivíduo 1 prefere uma mudança de x para y mais do que o Indivíduo 2 prefere uma mudança de x^* para y^* ” são consideradas sem sentido.

Em contextos de votação, essa suposição pode ser plausível, mas em contextos de avaliação de bem-estar – quando um planejador social busca ordenar diferentes alternativas sociais em uma ordem de bem-estar social – o uso de uma base informacional mais rica pode ser justificado. Sen (1970b) generalizou o modelo de Arrow para incorporar essa base mais ampla.

Como antes, considere um conjunto $N = \{1, 2, \dots, n\}$ de indivíduos ($n \geq 2$) e um conjunto $X = \{x, y, z, \dots\}$ de alternativas sociais. Agora, cada indivíduo $i \in N$ tem uma *função de bem-estar* W_i sobre essas alternativas, a qual atribui um número real $W_i(x)$ a cada alternativa $x \in X$ interpretado como uma medida do bem-estar de i sob a alternativa x . Qualquer função de bem-estar em X induz uma ordenação em X , mas o inverso não é verdadeiro: as funções de bem-estar codificam mais informações. Uma combinação de funções de bem-estar entre os indivíduos, $\langle W_1, W_2, \dots, W_n \rangle$, é chamada de perfil.

Um *funcional de bem-estar social* (FLBS), também denominado F , é uma função que atribui a cada perfil $\langle W_1, W_2, \dots, W_n \rangle$ (em algum domínio de perfis admissíveis) uma relação de preferência social $R = F(W_1, W_2, \dots, W_n)$ em X , com a interpretação usual. Novamente, quando F é claro no contexto, escrevemos R para a relação de preferência social correspondente a $\langle W_1, W_2, \dots, W_n \rangle$. O resultado de um FLBS é semelhante ao de uma regra de agregação de preferências (novamente, não incluímos a completude ou transitividade de R na definição³⁴), mas os dados iniciais são mais ricos.

O que ganhamos com isso depende de quanto dessa base informacional mais rica nos permitimos usar para determinar as preferências da sociedade: tecnicamente, depende de nossa suposição sobre mensurabilidade e comparabilidade interpessoal de bem-estar.

4.1 Mensurabilidade e comparabilidade interpessoal de bem-estar

Ao atribuir números reais às alternativas, os perfis de bem-estar contêm muitas informações além dos perfis de ordenações em X que eles induzem. Em particular, muitas atribuições diferentes de **números** para alternativas podem dar origem às mesmas **ordenações**. Contudo, podemos não considerar todas essas informações significativas. Parte disso pode ser um artefato da representação numérica. Por exemplo, a diferença entre o perfil $\langle W_1, W_2, \dots, W_n \rangle$ e sua versão ampliada $\langle 10*W_1, 10*W_2, \dots, 10*W_n \rangle$, onde tudo é o mesmo em termos proporcionais, pode ser como a diferença entre as medidas de comprimento em centímetros e em polegadas. Os dois perfis podem ser vistos como representações alternativas de exatamente as mesmas informações, apenas em escalas diferentes.

Para expressar diferentes suposições sobre quais informações são verdadeiramente codificadas por um perfil de funções de bem-estar e quais informações não são (e devem, portanto, ser vistas, na melhor das hipóteses, como um artefato da representação numérica), é útil introduzir a noção de *declarações significativas*. Alguns exemplos de declarações sobre o bem-estar individual que são candidatas a declarações significativas são os seguintes (LIST, 2003b; ver

³⁴ Sen, assim como Arrow em sua definição de funções de bem-estar social (enquanto opostas a funcionais), exigiu que R fosse uma ordenação por definição.

também BOSSERT; WEYMARK, 1996, seção 5):

Uma comparação de nível: o bem-estar do indivíduo i sob a alternativa x é pelo menos tão grande quanto o bem-estar do indivíduo j sob a alternativa y , formalmente $W_i(x) \geq W_j(y)$. (A comparação é *intrapessoal* se $i = j$, e *interpessoal* se $i \neq j$.)

Uma comparação de unidade: a proporção de [ganho ou perda de bem-estar do indivíduo i se mudarmos da alternativa y_1 para a alternativa x_1] para [ganho ou perda de bem-estar do indivíduo j se mudarmos da alternativa y_2 para a alternativa x_2] é λ , onde λ é algum número real, formalmente $(x_1 - y_1) / (x_2 - y_2) = \lambda$. (Novamente, a comparação é *intrapessoal* se $i = j$, e *interpessoal* se $i \neq j$.)

Uma comparação zero: o bem-estar do indivíduo i sob a alternativa x é maior que / igual a / menor que zero, formalmente $\text{sin}(\lambda(W_i(x))) = \lambda$, onde $\lambda \in \{-1, 0, 1\}$ e sin é uma função de valor real que mapeia estritamente números negativos para -1, zero para 0 e números estritamente positivos para +1.

A interpretação de Arrow, conforme mostrado, é que apenas as comparações no nível intrapessoal são significativas, enquanto todos os outros tipos de comparações não o são. Sen (1970b) formalizou várias suposições sobre mensurabilidade e comparabilidade interpessoal de bem-estar ao (i) definir uma relação de equivalência nos perfis de bem-estar que especifica quando dois perfis contam como “contendo as mesmas informações” e (ii) exigir quaisquer perfis na mesma classe de equivalência para gerar a mesma ordenação de preferência social. Dos três tipos de comparação introduzidos acima, as significativas são aquelas que são invariáveis em cada classe de equivalência. A suposição ordinalista de Arrow pode ser assim expressada:

Mensurabilidade ordinal sem comparabilidade interpessoal (OSC): Dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = \varphi_i(W_i)$, onde φ_i

é alguma transformação monotônica positiva, possivelmente diferente para diferentes indivíduos.

Assim, as funções de bem-estar individuais em qualquer perfil podem ser arbitrariamente transformadas de forma monotônica (“alongadas ou comprimidas”) sem perda de informação, descartando assim quaisquer comparações interpessoais ou mesmo comparações de unidades intrapessoais.

Se o bem-estar é cardinalmente mensurável, mas ainda interpessoalmente não comparável, temos:

Mensurabilidade cardinal sem comparabilidade interpessoal (CSC): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = a_i W_i + b_i$, onde a_i s e b_i s são números reais (com $a_i > 0$), possivelmente diferentes para indivíduos diferentes.

Aqui, a função de bem-estar de cada indivíduo é única até as transformações afins positivas (“escalonamento e deslocamento”), mas ainda não há uma escala comum entre os indivíduos. Isso torna as comparações de nível e unidade intrapessoais significativas, mas exclui as comparações interpessoais e as comparações zero.

A comparabilidade de nível interpessoal é alcançada sob a seguinte variante enriquecida de mensurabilidade ordinal:

Mensurabilidade ordinal com comparabilidade de nível interpessoal (OCN): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = \varphi_i(W_i)$, onde φ é a mesma transformação monotônica positiva para todos os indivíduos.

Aqui, um perfil de funções de bem-estar individual pode ser transformado arbitrariamente de forma monotônica (“alongado ou comprimido”) sem perda de informação, mas a mesma transformação deve ser usada para todos os indivíduos, tornando assim as comparações de nível interpessoal significativas.

A comparabilidade da unidade interpessoal é alcançada sob a seguinte variante enriquecida de mensurabilidade cardinal:

Mensurabilidade cardinal com comparabilidade de unidade interpessoal (CCU): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = aW_i + b_i$ onde a é o mesmo número real para todos os indivíduos ($a > 0$) e os b_i s são números reais.

Aqui, as funções de bem-estar em cada perfil podem ser reescaladas e deslocadas sem perda de informação, mas o mesmo múltiplo escalar (embora não necessariamente a mesma constante de deslocamento) deve ser usado para todos os indivíduos, tornando assim as comparações de unidades interpessoais significativas.

As comparações de zero, finalmente, tornam-se significativas sob a seguinte variante enriquecida de mensurabilidade ordinal (LIST, 2001):

Mensurabilidade ordinal com comparabilidade zero (OSC + 0): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = \varphi_i(W_i)$, onde φ_i é alguma transformação monotônica positiva e com preservação de zero, possivelmente diferente para diferentes indivíduos. (Aqui a preservação de zero significa que $\varphi_i(0) = 0$.)

Isso permite o alongamento e a compressão arbitrários das funções de bem-estar individual sem perda de informações, **desde que o nível de bem-estar de zero permaneça fixo**, garantindo assim a comparabilidade zero.

Vários outros pressupostos de mensurabilidade e comparabilidade interpessoal foram discutidos na literatura. O seguinte garante a significância das comparações interpessoais de ambos os níveis e unidades:

Mensurabilidade cardinal com comparabilidade interpessoal total (CCT): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = aW_i + b$ onde a, b são os mesmos números reais para todos os indivíduos ($a > 0$).

Por último, as comparações intra e interpessoais de todos os **três** tipos (nível, unidade e zero) são significativas se aceitarmos o seguinte:

Mensurabilidade da razão da escala com comparabilidade interpessoal total (RCT): dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ contêm a mesma informação sempre que, para cada $i \in N$, $W_i^* = aW_i$, onde a é o mesmo número real para todos os indivíduos ($a > 0$).

Qual pressuposto se justifica depende de como se interpreta o bem-estar. Se o bem-estar é uma **utilidade hedônica**, que pode ser experimentada apenas a partir de uma perspectiva de primeira pessoa, as comparações interpessoais são mais difíceis de justificar do que se o bem-estar é a **satisfação objetiva de preferências ou desejos subjetivos** (abordagem da *satisfação do desejo*) ou um **bem objetivo ou estado** (abordagem da *lista de objetivos*) (por exemplo HAUSMAN, 1995; LIST, 2003b). A abordagem da satisfação do desejo pode tornar as comparações interpessoais empiricamente significativas (ao relacionar os níveis máximos e mínimos interpessoalmente significativos de bem-estar para cada indivíduo com vistas a obtenção de suas alternativas mais e menos preferidas), mas às custas de enfrentar problemas de gostos caros ou de preferências adaptativas (HAUSMAN, 1995). Unidades de bem-estar baseadas em recursos, em funcionamentos ou em bens primários, por outro lado, podem permitir comparações interpessoais empiricamente significativas e menos moralmente problemáticas.

4.2 A possibilidade de agregação de bem-estar

Uma vez que introduzimos comparações interpessoais de níveis ou unidades de bem-estar, ou comparações zero, existem FLBSs possíveis que satisfazem os análogos às condições de Arrow, bem como desideratos mais fortes. Em um contexto de agregação de bem-estar, a impossibilidade de Arrow pode, portanto, ser atribuída a uma falta de comparabilidade interpessoal.

Conforme mencionado, um FLBS respeita uma determinada suposição sobre mensurabilidade e comparabilidade interpessoal se, para quaisquer dois

perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ que são considerados como contendo a mesma informação, temos $F(W_1, W_2, \dots, W_n) = F(W_1^*, W_2^*, \dots, W_n^*)$. As condições e o teorema de Arrow podem ser reformulados da seguinte forma:

Domínio universal: o domínio de F é o conjunto de todos os perfis logicamente possíveis de funções de bem-estar individuais.

Ordenação: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ no domínio de F , a relação de preferência social R é completa e transitiva.

Princípio fraco de Pareto: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ no domínio de F , se para todo $i \in N$ $W_i(x) > W_i(y)$, então xPy .

Independência de alternativas irrelevantes: para quaisquer dois perfis $\langle W_1, W_2, \dots, W_n \rangle$ e $\langle W_1^*, W_2^*, \dots, W_n^* \rangle$ no domínio de F e qualquer $x, y \in X$, se para todo $i \in N$ $W_i(x) = W_i^*(x)$ e $W_i(y) = W_i^*(y)$, então xRy se e somente se xR^*y .

Não-ditadura: não existe um indivíduo $i \in N$ tal que, para todos $\langle W_1, W_2, \dots, W_n \rangle$ no domínio de F e todos os $x, y \in X$, $W_i(x) > W_i(y)$ implica xPy .

Teorema: sob OSC, ou CSC, como Sen (1970b) demonstrou, se $|X| > 2$, não existe nenhum FLBS que satisfaça domínio universal, ordenação, princípio fraco de Pareto, independência de alternativas irrelevantes e não-ditadura.

Crucialmente, entretanto, cada um dos OCN, CCU e OSC + 0 é suficiente para a existência de FLBSs que satisfaçam todas as outras condições:

Teorema (combinando vários resultados da literatura, conforme ilustrado abaixo): sob cada um dos OCN, CCU e OSC + 0, existem FLBSs que satisfazem

domínio universal, ordenação, princípio fraco de Pareto, independência de alternativas irrelevantes e não-ditadura (bem como condições mais fortes).

Alguns exemplos de tais FLBSs vêm da filosofia política e da economia do bem-estar. Um possível FLBS sob OCN é uma versão do princípio da diferença de Rawls (1971).

Maximin: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $\min_{i \in N} (W_i(x)) \geq \min_{i \in N} (W_i(y))$.

Enquanto o *maximin* ordena as alternativas sociais em termos do nível de bem-estar do indivíduo em pior situação sozinho, sua extensão lexicográfica (*leximin*), que foi endossada pelo próprio Rawls, usa o nível de bem-estar do segundo indivíduo em pior situação como critério de desempate quando houver empate no nível do pior nível, o nível de bem-estar do terceiro indivíduo em pior situação como critério de desempate quando houver empate no segundo estágio, e assim por diante. (Importa observar, no entanto, que Rawls se concentrou em bens primários, em vez de bem-estar, como a unidade de análise relevante.) Isso satisfaz o *princípio forte de Pareto* (não apenas fraco), exigindo que se para todo $i \in N$ $W_i(x) \geq W_i(y)$, então xRy , e, além disso, se para algum $i \in N$ $W_i(x) > W_i(y)$, então xPy .

Um exemplo de um possível FLBS sob o CCU é o utilitarismo clássico.

Utilitarismo: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $W_1(x) + W_2(x) + \dots + W_n(x) \geq W_1(y) + W_2(y) + \dots + W_n(y)$.

Finalmente, um exemplo de um possível FLBS sob OSC + 0 é uma variante de uma medida de pobreza frequentemente usada, embora bastante simplista.

Uma regra de contagem de pessoas: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $|\{i \in N : W_i(x) < 0\}| < |\{i \in N : W_i(y) < 0\}|$ ou $|\{i \in N : W_i(x) < 0\}| = |\{i \in N : W_i(y) < 0\}|$ e xRy , onde $j \in N$ é algum indivíduo usado como critério de desempate previamente fixado.

Embora substancialmente menos convincentes do que as regras *maximin* ou utilitárias, as regras de contagem de pessoas exigem apenas comparabilidade zero do bem-estar (LIST, 2001).

Uma conclusão importante, portanto, é que o princípio da diferença de Rawls, o princípio utilitário clássico e até mesmo o método de contagem de pessoas para medição da pobreza podem ser vistos como soluções para o problema de agregação de Arrow que se tornam possíveis uma vez que transpomos o arcabouço de Arrow de preferências ordinais e interpessoalmente não-comparáveis.

Sob CCT, pode-se fornecer uma caracterização simultânea do *maximin* Rawlsiano e do utilitarismo (Deschamps e Gevers, 1978). Ela se utiliza de dois axiomas adicionais. Um, *equidade mínima*, requer “que uma pessoa que vai estar melhor de qualquer forma não precisa rigorosamente **sempre** fazer as coisas como deseja” (SEN, 1977, p. 1548), e outro, *separabilidade*, requer que dois perfis de bem-estar que coincidam para algum subconjunto $M \subseteq N$, enquanto todos em $N \setminus M$ são indiferentes entre todas as alternativas em X , resultem na mesma ordenação social.

Teorema (DESCHAMPS; GEVERS, 1978): sob o CCT, qualquer FLBS que satisfaça domínio universal, ordenação, princípio forte de Pareto, independência de alternativas irrelevantes, anonimato (conforme o teorema de May), equidade mínima e separabilidade é ou *leximin* ou de tipo *utilitário* (o que significa que, exceto possivelmente quando há empates na soma total de bem-estar, ela coincide com a FLBS utilitária definida acima).

Finalmente, as informações adicionais disponíveis sob RCT tornam os FLBSs “prioritários” possíveis³⁵. Como os FLBSs utilitários, eles ordenam as alternativas sociais com base nas somas de bem-estar entre os indivíduos em N , mas em vez de somar o bem-estar diretamente, eles somam o bem-estar transformado de forma côncava, dando maior peso marginal aos níveis mais baixos de bem-estar.

³⁵ Tecnicamente, isso requer uma restrição de domínio para perfis positivos de bem-estar.

Prioritarismo: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ e qualquer $x, y \in X$, xRy se e somente se $W_1^r(x) + W_2^r(x) + \dots + W_n^r(x) \geq W_1^r(y) + W_2^r(y) + \dots + W_n^r(y)$, onde $0 < r < 1$.

O prioritarismo requer RCT e não apenas CCT porque, dado seu desenho, a ordenação social prioritária para qualquer perfil de bem-estar não é invariável sob mudanças nos níveis de bem-estar (deslocamento).

4.3 Outros tópicos

A presente estrutura de agregação de bem-estar foi aplicada a várias outras áreas. Foi generalizado para problemas de escolha de população variável, de modo a formalizar a ética populacional na tradição de Parfit (1984). Aqui, devemos ordenar as alternativas sociais (p. ex., mundos possíveis) em que existem diferentes indivíduos. Assuma que $N(x)$ denota o conjunto de indivíduos existentes sob a alternativa x . Por exemplo, o conjunto $N(x)$ pode diferir do conjunto $N(y)$, quando x e y são alternativas distintas (isso generaliza nossa suposição anterior de um conjunto fixo N). O caso da população variável levanta questões como se um mundo com número menor de indivíduos em melhor situação é melhor, igualmente bom ou pior do que um mundo com um número maior de indivíduos em situação pior. (O foco aqui está em questões axiológicas sobre a bondade relativa de tais mundos, não questões normativas sobre a correção ou torpeza em produzi-los).

Parfit (1984) e outros argumentaram que o utilitarismo clássico está sujeito à *conclusão repugnante*: um mundo com grande número de indivíduos cujos níveis de bem-estar estão pouco acima de zero poderia ter uma soma total de bem-estar maior e, portanto, contar como melhor do que um mundo com número menor de indivíduos muito abonados.

Blackorby, Donaldson e Bossert (2005) caracterizaram axiomáticamente diferentes métodos de agregação de bem-estar de população variável que evitam a conclusão repugnante e satisfazem alguns outros desideratos. Uma solução é a seguinte:

Utilitarismo de nível crítico: para qualquer perfil $\langle W_1, W_2, \dots, W_n \rangle$ e qualquer $x, y \in X$, xRy se e somente

se $\sum_{i \in N(x)} [W_i(x) - c] \geq \sum_{i \in N(y)} [W_i(y) - c]$, onde $c \geq 0$ é algum “nível crítico” de bem-estar acima do qual a qualidade de vida conta como “decente / boa”.

O utilitarismo de nível crítico evita a conclusão repugnante quando o parâmetro c é definido de forma grande o suficiente. Ele requer uma mensurabilidade de bem-estar mais forte do que o utilitarismo clássico, uma vez que gera uma ordenação social R que geralmente não é invariante sob o reescalonamento das unidades de bem-estar ou mudanças nos níveis de bem-estar. Mesmo a estrutura rica da RCT forçaria o nível crítico c a ser zero, colapsando o utilitarismo de nível crítico no utilitarismo clássico e tornando-o vulnerável à conclusão repugnante novamente. Conforme Blackorby, Bossert e Donaldson (1999, p. 420) observam,

[A]lguns ambientes de informação que são eticamente adequados em ambientes de população fixa têm consequências eticamente pouco atraentes em ambientes de população variável.

Assim, no caso da população variável, um afastamento mais significativo da estrutura informacional limitada do modelo original de Arrow é necessário para evitar resultados impossíveis.

A abordagem FLBS foi generalizada para o caso em que cada indivíduo tem várias funções de bem-estar (por exemplo, um k -tuplo delas), capturando (i) múltiplas opiniões sobre o bem-estar de cada indivíduo (por exemplo ROBERTS, 1995; OOGHE; LAUWERS, 2005) ou (ii) múltiplas dimensões de bem-estar (por exemplo LIST, 2004a). Nesse caso, nos deparamos não apenas com questões de mensurabilidade e comparabilidade interpessoal, mas também com questões de comparabilidade inter-opinião ou interdimensional. Para obter resultados de possibilidade convincentes, é necessária a comparabilidade entre os indivíduos e as dimensões/opiniões. Uma literatura relacionada aborda a medição da desigualdade multidimensional. Para uma revisão introdutória, consulte Weymark (2006).

Finalmente, na filosofia da biologia, as estruturas FLBS unidimensionais e multidimensionais foram usadas por Okasha (2009) e Bossert, Qi e Weymark (2013) para analisar a noção de aptidão de grupo, definida como uma função de indicadores de aptidão individual.

5. Agregação de julgamentos

Um ramo mais recente da teoria da escolha social é a teoria da agregação de julgamentos. Seu desenvolvimento é motivado pela observação de que votos, ordenações ou funções de bem-estar em relação a várias alternativas não são os únicos objetos que podemos desejar agregar de um nível individual a um nível coletivo. Muitos órgãos de tomada de decisão, como legislaturas, tribunais colegiados, painéis de especialistas e outros comitês, enfrentam um “agregando” mais complexo. Em particular, eles podem ter que agregar conjuntos individuais de julgamentos sobre proposições múltiplas e logicamente conectadas em conjuntos coletivos de julgamentos.

Um tribunal, para julgar se um réu é responsável por violação de contrato, deve analisar se o contrário era válido e se houve uma violação. Um painel de especialistas pode ter que julgar se as concentrações atmosféricas de gases do efeito estufa excederão um determinado limite em 2050, se há uma cadeia causal de maiores concentrações de gases do efeito estufa ao aumento da temperatura e se a temperatura aumentará. Legisladores podem ter de julgar se um fim específico é socialmente desejável, se uma política pública proposta é o melhor meio para atingir esse fim e se a política deve ser adotada.

Esses problemas não podem ser formalizados em modelos padrão de agregação de preferências, pois os objetos da agregação não são ordenações, mas conjuntos de julgamentos sobre proposições múltiplas. A teoria da agregação de julgamentos representa esses agregandos na lógica proposicional (ou outra lógica adequada). O campo foi inspirado pelo “paradoxo doutrinário” da teoria do direito, com o qual começamos.

5.1 O “paradoxo doutrinário” e o “dilema discursivo”

Kornhauser e Sager (1986) descreveram o seguinte problema. (Um problema estruturalmente semelhante foi descoberto por Vacca, 1921 e, como Elster, 2013 aponta, por Poisson, 1837). Um colegiado de três juízes em um tribunal deve fazer julgamentos sobre as seguintes proposições:

- p : o réu foi contratualmente obrigado a não fazer a ação X .
- q : o réu executou a ação X .
- r : o réu é responsável por quebrar o contrato.

De acordo com a doutrina jurídica, as **premissas** p e q são, em conjunto, necessárias e suficientes para a conclusão r . Suponha que os juízes individuais tenham as opiniões mostradas na Tabela 5.

	p (obrigação)	q (ação)	r (responsabilidade)
Juiz 1	Verdadeiro	Verdadeiro	Verdadeiro
Juiz 2	Falso	Verdadeiro	Falso
Juiz 3	Verdadeiro	Falso	Falso
Maioria	Verdadeiro	Verdadeiro	Falso

Tabela 5: Um exemplo do “paradoxo doutrinário”

Embora cada juiz individual respeite a doutrina jurídica relevante, há uma maioria a favor de p , uma maioria a favor de q e, ainda assim, uma maioria contra r – em violação à doutrina jurídica. O tribunal enfrenta um dilema: ele pode *tanto* seguir os julgamentos da maioria nas premissas (p e q) e chegar a um veredicto “responsável” por inferência lógica (a abordagem *questão por questão* ou *baseada em premissa*); ou adotar o julgamento da maioria sobre a conclusão (r) e chegar a um veredicto “não responsável”, ignorando os julgamentos da maioria sobre as premissas (a abordagem *caso a caso* ou *baseada em conclusões*). O “paradoxo doutrinário” de Kornhauser e Sager consiste no fato de que essas duas abordagens podem levar a resultados opostos.

Podemos aprender outra lição com este exemplo. Em relação à doutrina jurídica, os julgamentos da maioria são logicamente inconsistentes. Expresso formalmente, o conjunto de proposições aceitas pela maioria, $\{p, q, \text{não } r\}$, é inconsistente em relação à restrição r **se e somente se** (p e q). Essa observação foi o ponto de partida da literatura mais recente, baseada em lógica formal, sobre agregação de julgamentos (começando com um modelo e resultado de impossibilidade em List e Pettit, de 2002).

A possibilidade de julgamentos por maioria inconsistentes não está ligada à presença de uma doutrina jurídica ou de outra restrição secundária explícita (conforme apontado por Pettit, em 2001, que chamou esse fenômeno de “dilema discursivo”). Suponha, por exemplo, que um painel de especialistas tenha que fazer julgamentos sobre três proposições (e suas negações):

- p : o CO_2 atmosférico excederá 600ppm em 2050.
- *se p então q* : se o CO_2 atmosférico exceder esse nível até 2050, haverá um aumento de temperatura de mais de $3,5^\circ$ até 2010.

- q : haverá um aumento de temperatura de mais de 3,5° até 2010.

Se os julgamentos individuais são como mostrados na Tabela 6, os julgamentos da maioria são inconsistentes: apesar dos julgamentos individualmente consistentes, o conjunto de proposições aceitas pela maioria, $\{p, \text{se } p, \text{então } q, \text{não } q\}$, é logicamente inconsistente.

	p	$\text{se } p \text{ então } q$	q
<i>Especialista 1</i>	Verdadeiro	Verdadeiro	Verdadeiro
<i>Especialista 2</i>	Falso	Verdadeiro	Falso
<i>Especialista 3</i>	Verdadeiro	Falso	Falso
<i>Maioria</i>	Verdadeiro	Verdadeiro	Falso

Tabela 6: Uma inconsistência majoritária

5.2 As condições para julgamentos majoritários inconsistentes

Os padrões de julgamentos nas Tabelas 5 e 6 são estruturalmente equivalentes ao padrão de preferências que levam ao paradoxo de Condorcet, quando reinterpretemos essas preferências como julgamentos sobre proposições da forma “ x é preferível a y ”, “ y é preferível a z ”, e assim por diante, conforme mostrado na Tabela 7 (LIST; PETTIT, 2004). Uma interpretação anterior das preferências nessa linha pode ser encontrada em Guilbaud, de 1952 (1966). Aqui, o conjunto de proposições aceitas pela maioria é inconsistente em relação à restrição de transitividade.

	$\text{“}x \text{ é preferível a } y\text{”}$	$\text{“}y \text{ é preferível a } z\text{”}$	$\text{“}x \text{ é preferível a } z\text{”}$
<i>Indivíduo 1</i> (prefere x a y a z)	Verdadeiro	Verdadeiro	Verdadeiro
<i>Indivíduo 2</i> (prefere y a z a x)	Falso	Verdadeiro	Falso
<i>Indivíduo 3</i> (prefere z a x a y)	Verdadeiro	Falso	Falso
<i>Maioria</i> (prefere x a y a z a x , um “ciclo”)	Verdadeiro	Verdadeiro	Falso

Tabela 7: Paradoxo de Condorcet, reinterpretado proposicionalmente.

Um resultado combinatório geral inclui todos esses fenômenos. Chamamos um conjunto de proposições **minimamente inconsistente** se for um conjunto logicamente inconsistente, mas todos os seus subconjuntos são consistentes.

Proposição (DIETRICH; LIST, 2007a; NEHRING; PUPPE, 2007): a votação majoritária por proposições pode gerar julgamentos coletivos inconsistentes se e somente se o conjunto de proposições (e suas negações) sobre as quais se realizam os julgamentos tem um subconjunto minimamente inconsistente de três ou mais proposições.

Nos exemplos das Tabelas 6, 5 e 7, os conjuntos minimamente inconsistentes relevantes de tamanho (pelo menos) três são: $\{p; \text{se } p, \text{ então } q; \text{ não } q\}$, que é minimamente inconsistente *simpliciter*; $\{p, q, \text{ não-}r\}$, que é minimamente inconsistente em relação à restrição lateral r se e somente se $(p \text{ e } q)$; e $\{“x \text{ é preferível a } y”, “y \text{ é preferível a } z”, “z \text{ é preferível a } x”\}$, o que é minimamente inconsistente em relação a uma restrição de transitividade na preferibilidade.

5.3 Um modelo geral e um resultado de impossibilidade simples

O modelo básico de agregação de julgamentos pode ser definido no formato a seguir (LIST; PETTIT, 2002). Seja $N = \{1, 2, \dots, n\}$ um conjunto de indivíduos ($n \geq 2$). As proposições sobre as quais os julgamentos devem ser feitos são representadas por sentenças da lógica proposicional (ou alguma outra lógica expressivamente mais rica, como a de predicado, a modal ou a condicional; ver Dietrich, 2007). Definimos a *agenda*, X , como um conjunto finito de proposições, fechado sob uma única negação³⁶. Por exemplo, X poderia ser $\{p, \neg p, p \rightarrow q, \neg(p \rightarrow q), q, \neg q\}$, como no caso do painel de especialistas.

³⁶ Formalmente, $X = \{p, \neg p: p \in X^+\}$, onde X^+ é um conjunto de proposições não-negadas. Para evitar tecnicidades, assumimos que X não contém proposições contraditórias ou tautológicas.

Cada indivíduo $i \in N$ tem um **conjunto de julgamentos** J_i , definido como um subconjunto $J_i \subseteq X$ e interpretado como o conjunto de proposições que o indivíduo i aceita. Um conjunto de julgamentos é **consistente** se for um conjunto de proposições logicamente consistente³⁷ e **completo** (em relação a X) se contiver um membro de cada par proposição-negação $p, \neg p \in X$.

Uma combinação de conjuntos de julgamento entre os indivíduos, $\langle J_1, J_2, \dots, J_n \rangle$, é chamada de *perfil*. Uma *regra de agregação de julgamentos*, F , é uma função que atribui a cada perfil $\langle J_1, J_2, \dots, J_n \rangle$ (em algum domínio de perfis admissíveis) um conjunto de julgamento coletivo $J = F(J_1, J_2, \dots, J_n) \subseteq X$, interpretado como o conjunto de proposições aceitas pelo grupo como um todo. Como antes, quando F é claro no contexto, escrevemos J para o conjunto de julgamento coletivo correspondente a $\langle J_1, J_2, \dots, J_n \rangle$. Novamente, por generalidade, não desenvolvemos nenhum requisito de racionalidade em J (como consistência ou completude) na definição de uma regra de agregação de julgamentos.

O exemplo mais simples de uma regra de agregação de julgamentos é a *votação majoritária por proposição*. Aqui, para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$, $J = \{p \in X : |\{i \in N : p \in J_i\}| > n/2\}$. Como vimos, isso pode produzir julgamentos coletivos inconsistentes.

Considere as seguintes condições em uma regra de agregação:

Domínio universal: o domínio de F é o conjunto de todos os perfis logicamente possíveis de conjuntos de julgamento individual consistentes e completos.

Racionalidade coletiva: para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$ no domínio de F , o conjunto de julgamento coletivo J é consistente e completo.

Anonimato: para quaisquer dois perfis $\langle J_1, J_2, \dots, J_n \rangle$ e $\langle J_1^*, J_2^*, \dots, J_n^* \rangle$ que são permutações entre si, $F(J_1, J_2, \dots, J_n) = F(J_1^*, J_2^*, \dots, J_n^*)$

³⁷ Em princípio, consistência pode ser definida em relação a alguma restrição lateral, tal como a doutrina jurídica no exemplo do “paradoxo doutrinário”.

Sistematicidade: para quaisquer dois perfis $\langle J_1, J_2, \dots, J_n \rangle$ e $\langle J_1^*, J_2^*, \dots, J_n^* \rangle$ no domínio de F e qualquer $p, q \in X$, se para todo $i \in N$ $p \in J_i$ se e somente se $q \in J_i^*$, então $p \in J$ se e somente se $q \in J^*$.

As três primeiras condições são análogas a domínio universal, ordenação e anonimato na agregação de preferências. A última é a contrapartida da independência de alternativas irrelevantes, embora mais forte: requer que (i) os julgamentos coletivos sobre qualquer proposição $p \in X$ (da qual uma proposição de ordenação binária como “ x é preferível a y ” é um caso especial) dependem apenas de proposições individuais em p (a parte da **independência**) e (ii) o padrão de dependência entre os julgamentos individuais e coletivos seja o mesmo em todas as proposições em X (a parte da **neutralidade**). Formalmente, a **independência** é o caso especial com a quantificação restrita a $p = q$. A votação majoritária por proposição satisfaz todas essas condições, exceto a parte de consistência da racionalidade coletiva.

Teorema (List e Pettit, 2002): se $\{p, q, p \wedge q\} \subseteq X$ (onde p e q são proposições mutuamente independentes e “ $\wedge$ ” também pode ser substituído por “ $\vee$ ” ou “ $\rightarrow$ ”), não existe regra de agregação de julgamentos que satisfaça domínio universal, racionalidade coletiva, anonimato e sistematicidade.

Como outros teoremas de impossibilidade, este resultado é melhor interpretado como descrevendo as compensações (*trade-offs*) entre diferentes condições em uma regra de agregação. O resultado foi generalizado e fortalecido de várias maneiras, começando com a prova de Pauly e van Hees (2006) de que a impossibilidade persiste se o anonimato for enfraquecido para não-ditadura. Para outras generalizações, confira Dietrich (2006) e Mongin (2008).

5.4 Teoremas de impossibilidade e possibilidade mais gerais

Conforme vimos, na agregação de preferência, o “limite” entre os resultados de possibilidade e impossibilidade é fácil de traçar: quando há apenas duas alternativas de decisão, todos os desideratos em uma regra de agregação de preferência

revisados acima podem ser satisfeitos (e a regra da maioria realiza essa função); quando há três ou mais alternativas, há resultados de impossibilidade. Na agregação de julgamentos, ao contrário, o quadro é mais complicado. O que importa não é o número de proposições em X , mas a natureza das interconexões lógicas entre elas.

Os resultados de impossibilidade na agregação de julgamentos têm a seguinte forma genérica: para uma determinada classe de agendas, as regras de agregação que satisfazem um conjunto particular de condições (geralmente, uma condição de domínio, uma condição de racionalidade e algumas condições de responsividade) são inexistentes ou degeneradas (p. ex., ditatorial). Diferentes tipos de agendas acionam diferentes instâncias deste esquema, com condições mais fortes ou mais fracas impostas na regra de agregação dependendo das propriedades dessas agendas. Para uma revisão mais detalhada, consulte List (2012). A significância das propriedades combinatórias da agenda foi descoberta primeiro por Nehring e Puppe (2002) em um arcabouço matematicamente relacionado, mas com interpretação distinta (escolha social resistente à estratégia sobre os chamados espaços de propriedade). Três tipos de agenda se destacam:

Uma agenda não-simples: X tem um subconjunto minimamente inconsistente de três ou mais proposições.

Uma agenda par-negável: X tem um subconjunto minimamente inconsistente Y que pode ser tomado consistente pela negação de um par de proposições contidos nele. De forma equivalente, X não é isomórfico a um conjunto de proposições cujos únicos conectivos são $\neg$ e $\leftrightarrow$. (DOKOW; HOLZMAN, 2010a.)

Uma agenda trajetória-conexa, ou *totalmente bloqueada*, como em Nehring e Puppe (2002): para qualquer $p, q \in X$, há uma sequência $p_1, p_2, \dots, p_k \in X$ com $p_1 = p$ e $p_k = q$ tal que p_1 condicionalmente acarreta p_2 , p_2 acarreta condicionalmente p_3 , ..., e p_{k-1} acarreta condicionalmente p_k . (Aqui, p_i condicionalmente acarreta p_j se $p_i \cup Y$ acarreta p_j para algum $Y \subseteq X$ consistente com p_i e $\neg p_j$.)

Algumas agendas têm duas ou mais dessas propriedades. As agendas em nossos exemplos de “paradoxo doutrinário” e “dilema discursivo” são não-simples e par-negáveis. A *agenda de preferências*, $X = \{“x \text{ é preferível a } y”, “y \text{ é preferível a } x”, “x \text{ é}$

preferível a z ", " z é preferível a x ", ...}, é não-simples, par-negável e trajetória-conectada (assumindo que a preferência é transitiva e completa). O seguinte resultado é consistente:

Teorema (DIETRICH; LIST, 2007b; DOKOW; HOLZMAN, 2010a), desenvolvido a partir de Nehring e Puppe (2002): se X for não-simples, par-negável e trajetória-conectada, não existe regra de agregação de julgamentos que satisfaça domínio universal, racionalidade coletiva, independência, preservação da unanimidade (exigindo que, para qualquer perfil unânime $\langle J, J, \dots, J \rangle$, $F(J, J, \dots, J) = J$), e não-ditadura³⁸.

Aplicado à agenda de preferências, este resultado produz o teorema de Arrow (para ordenações de preferência estritas) como um corolário. Predecessores desse resultado podem ser encontrados em List e Pettit (2004) e em Nehring (2003).³⁹ Assim, a agregação de preferências arrowiana pode ser reinterpretada como um caso especial de agregação de julgamentos.

A literatura contém diversas variantes deste teorema. Uma variante descarta a propriedade de trajeto-conexidade da agenda e fortalece a independência para a sistematicidade. Uma segunda variante descarta a propriedade de par-negabilidade da agenda e impõe uma condição de monotonicidade na regra de agregação (exigindo que o suporte adicional nunca prejudique uma proposição aceita) (NEHRING; PUPPE, 2010). O último resultado foi provado pela primeira vez na supracitada estrutura matematicamente relacionada de Nehring e Puppe (2002). Uma variante final descarta tanto a propriedade de trajeto-conexidade quanto de par-negabilidade, ao mesmo tempo que impõe sistematicidade e monotonicidade (*ibid.*).

Em cada caso, as propriedades da agenda não são apenas suficientes, mas também (se $n \geq 3$) necessárias para o resultado (NEHRING; PUPPE, 2002, 2010; DOKOW; HOLZMAN, 2010a). Observe também que a propriedade trajeto-conexidade implica não-simplicidade. Portanto, a não-simplicidade não precisa ser listada entre

³⁸ Veja também a observação sobre a relação entre as propriedades de trajeto-conexidade e de não-simplicidade no final desta subseção.

³⁹ Uma contribuição anterior matematicamente relacionada, embora distinta em sua interpretação, é o trabalho de Wilson sobre agregação abstrata (1975).

as condições do teorema, embora seja necessária nas variantes que eliminam a trajeto-conexidade.

5.5 Regras de agregação de julgamentos não-ditatoriais

5.5.1 Relaxando o domínio universal

Como na agregação de preferências, uma maneira de evitar os resultados de impossibilidade presentes é relaxar o domínio universal. Se o domínio dos perfis admissíveis de conjuntos de julgamento individual é restrito àqueles que satisfazem condições específicas de “coesão”, a votação majoritária por proposição produz julgamentos coletivos consistentes.

A condição de coesão mais simples é o *alinhamento unidimensional* (List, 2003c). Um perfil $\langle J_1, J_2, \dots, J_n \rangle$ é unidimensionalmente alinhado se os indivíduos em N podem ser ordenados da esquerda para a direita (por exemplo, em alguma dimensão cognitiva ou ideológica) tal que, para cada proposição $p \in X$, os indivíduos que aceitem p (ou seja, aqueles com $p \in J_i$) estão todos ou à esquerda ou à direita daqueles que rejeitam p (ou seja, aqueles com $p \notin J_i$), conforme ilustrado na Tabela 8. Para qualquer perfil desse tipo, os julgamentos majoritários são consistentes: o conjunto de julgamento do indivíduo mediano em relação à ordenação esquerda-direita prevalecerá (onde n é ímpar). Esse conjunto de julgamentos herdará sua consistência do indivíduo mediano, assumindo que os julgamentos individuais sejam consistentes. Por implicação, em domínios alinhados unidimensionalmente, a votação majoritária por proposição satisfará as demais condições das regras de agregação de julgamentos revisadas acima.

	Indivíduo 1	Indivíduo 2	Indivíduo 3	Indivíduo 4	Indivíduo 5
p	Verdadeiro	Verdadeiro	Falso	Falso	Falso
q	Verdadeiro	Verdadeiro	Verdadeiro	Verdadeiro	Falso
r	Falso	Falso	Falso	Verdadeiro	Verdadeiro
$p \wedge q \wedge r$	Falso	Falso	Falso	Falso	Falso

Tabela 8: Alinhamento unidimensional

Em analogia com o caso do pico único na agregação de preferências, várias condições menos restritivas já são suficientes para julgamentos majoritários

consistentes. Uma dessas condições (introduzida em Dietrich e List, em 2010a, em que uma revisão da literatura é apresentada) generaliza **restrição de valor por trios** de Sen. Um perfil $\langle J_1, J_2, \dots, J_n \rangle$ tem restrição de valor se todo subconjunto minimamente inconsistente $Y \subseteq X$ tem um par de elementos p, q tal que nenhum indivíduo $i \in N$ tem $\{p, q\} \subseteq J_i$. A restrição de valor evita que qualquer subconjunto minimamente inconsistente de X se torne aceito pela maioria e, portanto, garante julgamentos por maioria consistentes. Aplicada à agenda de preferências, a restrição de valor reduz-se à condição igualmente nomeada de Sen.

5.2.2 Relaxando a racionalidade coletiva

Enquanto o requisito de que os julgamentos coletivos sejam consistentes é amplamente aceito, o requisito de que os julgamentos coletivos sejam completos (em X) é fonte de debates. Em apoio à completude, pode-se dizer que uma dada proposição não seria incluída em X a menos que fosse ser julgada coletivamente. Contra a completude, pode-se dizer que há circunstâncias nas quais o nível de discordância sobre uma proposição particular (ou conjunto de proposições) é tão grande que formar uma visão coletiva sobre ela é indesejável ou contraproducente. Vários artigos oferecem resultados sobre a possibilidade ou a impossibilidade do relaxamento da completude (por exemplo

LIST; PETTIT, 2002; GÄRDENFORS, 2006; DIETRICH; LIST, 2007a, 2008; DOKOW; HOLZMAN, 2010b).

As regras de agregação de julgamentos que violam a completude coletiva mesmo satisfazendo (todas ou quase todas) as outras condições introduzidas acima incluem: *regra de unanimidade*, onde, para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$, $J = \{p \in X : p \in J_i \text{ para todos } i \in N\}$; *regras da supermaioria*, onde, para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$, $J = \{p \in X : |\{i \in N : p \in J_i\}| > qn\}$ para uma cota de aceitação adequada $q \in (0.5, 1)$; e *regras baseadas em conclusões*, onde um subconjunto $Y \subseteq X$ de proposições logicamente independentes (e suas negações) é designado como um conjunto de conclusões e $J = \{p \in Y : |\{i \in N : p \in J_i\}| > n/2\}$. No exemplo de tribunal com vários membros da Tabela 5, o conjunto de conclusões é simplesmente $Y = \{r, \neg r\}$.

Dado os conjuntos consistentes de julgamento individual, a regra da unanimidade garante conjuntos consistentes de julgamento coletivo, porque a interseção de vários conjuntos consistentes de proposições é sempre consistente. As regras da supermaioria também garantem conjuntos de julgamento coletivo

consistentes, desde que a cota q seja escolhida para ser pelo menos $(k - 1)/k$, onde k é o tamanho do maior subconjunto minimamente inconsistente de X . A razão é combinatória: quaisquer k supermaiorias distintas do tamanho relevante sempre terão pelo menos um indivíduo em comum. Assim, para que qualquer conjunto minimamente inconsistente de proposições (que é no máximo de tamanho k) seja aceito pela maioria, pelo menos um indivíduo teria que aceitar todas as proposições no conjunto, contradizendo a consistência desse indivíduo (DIETRICH; LIST, 2007a; LIST; PETTIT, 2002). Regras baseadas em conclusões, finalmente, produzem conjuntos consistentes de julgamento coletivo por construção, mas sempre deixam as não-conclusões sem decisão.

Gärdenfors (2006) e, de forma mais geral, Dietrich e List (2008) e Dokow e Holzman (2010b) mostraram que se – enquanto relaxamos a completude – exigimos que os conjuntos de julgamento coletivo sejam **dedutivamente fechados** (ou seja, para qualquer $p \in X$ acarretado por J , deve ser que $p \in J$), nos deparamos com um resultado de impossibilidade novamente. Para as mesmas agendas que levam ao resultado de impossibilidade revisado na Seção 5.4, não existe regra de agregação de julgamentos que satisfaça domínio universal, consistência coletiva e fechamento dedutivo, independência, preservação de unanimidade e não-oligarquia. Uma regra de agregação é chamada de *oligárquica* se houver um subconjunto previamente fixo $M \subseteq N$ (os “oligarcas”) tal que, para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$, $J = \{p \in X : p \in J_i \text{ para todos } i \in M\}$. A regra de unanimidade e as ditaduras são casos especiais com $M = N$ e $M = \{i\}$ para algum $i \in N$, respectivamente.

A desvantagem das regras de agregação oligárquica é que elas recaem na ditadura ou levam ao impasse com as menores divergências entre os oligarcas resultando em indecisão (já que todo oligarca tem poder de veto em todas as proposições).

5.5.3 Relaxando sistematicidade / independência

Uma variedade de regras de agregação de julgamentos torna-se possível quando relaxamos a sistematicidade / independência. Lembre-se de que a sistematicidade combina requisitos de independência e neutralidade. Relaxar apenas a neutralidade não nos leva muito longe, dado que para muitas agendas há resultados de impossibilidade apenas com independência, conforme ilustrado na Seção 5.4.

Uma classe muito discutida de regras de agregação que violam a independência é fornecida pelas *regras baseadas em premissas*. Aqui, um subconjunto

$Y \subseteq X$ de proposições logicamente independentes (e suas negações) é designado como um conjunto de *premissas*, como no exemplo do tribunal. Para qualquer perfil $\langle J_1, J_2, \dots, J_n \rangle$, $J = \{p \in X : J_Y \text{ implica } p\}$ onde J_Y são as proposições aceitas pela maioria entre as premissas, formalmente $\{p \in Y : |\{i \in N : p \in J_i\}| > n/2\}$. Informalmente, os votos da maioria são apresentados nas premissas, e os julgamentos coletivos sobre todas as outras proposições são determinados por implicação lógica. Se as premissas constituem uma base lógica para toda a agenda, uma regra baseada em premissas garante conjuntos de julgamento coletivo consistentes e (sem empates) completos. A presente definição segue List e Pettit (2002); para generalizações, consulte Dietrich e Mongin (2010). As propriedades procedimentais e epistêmicas das regras baseadas em premissas são discutidas em Pettit (2001), Chapman (2002), Bovens e Rabinowicz (2006), Dietrich (2006) e List (2006).

Uma generalização é oferecida pelas *regras de prioridade sequencial* (LIST, 2004b; DIETRICH; LIST, 2007a). Aqui, para cada perfil $\langle J_1, J_2, \dots, J_n \rangle$, as proposições em X são julgadas coletivamente em uma ordem fixa de prioridade, por exemplo, temporal ou epistêmica. O julgamento coletivo de cada proposição $p \in X$ é realizado da seguinte maneira. Se o julgamento majoritário em p for consistente com os julgamentos coletivos sobre proposições anteriores, este julgamento da maioria prevalece; caso contrário, o julgamento coletivo em p é determinado pelas implicações de julgamentos anteriores. Por construção, isso garante julgamentos coletivos consistentes e (sem empates) completos. No entanto, é **dependente de trajetória** (*path-dependent*): a ordem na qual as proposições são consideradas pode afetar o resultado, especificamente quando os julgamentos majoritários subjacentes são inconsistentes. Por exemplo, quando essa regra de agregação é aplicada aos perfis das Tabelas 5, 6 e 7 (mas não 8), os julgamentos coletivos dependem da ordem na qual as proposições são consideradas. Assim, as regras de prioridade sequencial são vulneráveis à **manipulação da agenda**. Fenômenos semelhantes ocorrem na votação majoritária em pares sequencial na agregação de preferências (por exemplo RIKER, 1982).

Outra classe proeminente de regras de agregação, que violam a independência, é oferecida pelas regras baseadas na distância, proposta em Pigozzi (2006), com base em Konieczny e Pino Pérez (2002). Confirma, também, Miller e Osherson (2009). Uma regra baseada na distância é definida em termos de alguma métrica de distância entre conjuntos de julgamento, por exemplo a *distância de Hamming*, onde, para quaisquer dois conjuntos de julgamento $J, J' \subseteq X$, $d(J, J') = |\{p \in X : \text{não } [p \in J \Leftrightarrow p \in J']\}|$. Cada perfil $\langle J_1, J_2, \dots, J_n \rangle$ é mapeado para um conjunto de julgamento

consistente e completo J que minimiza a distância da soma-total de cada um dos J s. As regras baseadas na distância podem ser interpretadas como capturando a ideia de identificar julgamentos de compromisso. Diferentemente das regras de prioridade baseadas em premissas ou sequenciais, as regras baseadas na distância não requerem uma distinção entre premissas e conclusões ou qualquer outra ordem de prioridade entre as proposições.

Como na agregação de preferências, o custo de relaxar a independência é a perda da resistência à estratégia. A conjunção de independência e monotonicidade é necessária e suficiente para a não manipulabilidade de uma regra de agregação de julgamentos por votação estratégica (DIETRICH; LIST, 2007c). Para resultados relacionados, consulte Nehring e Puppe (2002). Assim, geralmente não podemos alcançar a resistência à estratégia sem relaxar ou o domínio universal, ou a racionalidade coletiva, ou a preservação da unanimidade, ou a não-ditadura. Na prática, devemos, portanto, buscar maneiras de tornar as oportunidades de manipulação estratégica menos ameaçadoras.

6. Outros tópicos

Como deveria ser evidente, a teoria da escolha social é um campo vasto. As áreas não cobertas neste verbete, ou mencionadas apenas de passagem, incluem: teorias de divisão justa (como dividir um ou vários bens divisíveis ou indivisíveis, como bolos ou casas, entre vários requerentes). Confira, por exemplo, os trabalhos de Brams e Taylor (1996) e Moulin (2004). Sobre a teoria da escolha social comportamental (análise de evidências empíricas de comportamento de voto sob várias regras de agregação), confira Regenwetter *et al.* (2006), List, Luskin, Fishkin e McLean (2013). Acerca da teoria da escolha social empírica (análise de *surveys* e experimentos sobre as intuições das pessoas sobre justiça distributiva), consulte Gaertner e Schokkaert (2012). Em relação à teoria da escolha social computacional (análise de propriedades computacionais de regras de agregação, incluindo sua complexidade computacional), veja Bartholdi, Tovey e Trick (1989), Brandt, Conitzer e Endriss (2013). Sobre as teorias de agregação de probabilidade (que estudam a agregação de funções de probabilidade ou credibilidade), consulte Lehrer e Wagner (1981), McConway (1981), Genest e Zidek (1986), Mongin (1995), Dietrich e List (2007d); Por sua vez, sobre teorias de agregação de atitude geral (as que generalizam agregação de julgamentos de dois valores, agregação de

probabilidade/credibilidade e agregação de preferências), confira Dietrich e List (2010b), Dokow e Holzman (2010c). Acerca de estudo da tomada de decisão coletiva em animais não humanos (ou seja, estudo de decisões de grupo em uma variedade de espécies animais de insetos sociais a primatas), confira a pesquisa de Conradt e List (2009) e a edição especial que o texto introduz. E, por fim, em relação às aplicações à epistemologia social (a análise de estados doxásticos de grupo e sua relação com estados doxásticos individuais), veja Goldman (2004, 2010).

Referência bibliográfica

- ARROW, K. **Social Choice and Individual Values**. New York: Wiley, 1951/1963.
- AUSTEN-SMITH, D.; BANKS, J. S. Information Aggregation, Rationality, and the Condorcet Jury Theorem. **American Political Science Review**, v. 90, p. 34-45, 1996.
- BARTHOLDI, J. J.; TOVEY, C. A.; TRICK, M. A. The computational difficulty of manipulating an election. **Social Choice and Welfare**, v. 6, p. 227-241, 1989.
- BEN-YASHAR, R.; NITZAN, S. The optimal decision rule for fixed-size committees in dichotomous choice situations: the general result. **International Economic Review**, v. 38, p. 175-186, 1997.
- BEREND, D.; PAROUSH, J. When is Condorcet's Jury Theorem valid? **Social Choice and Welfare**. v. 15, p. 481-488, 1998.
- BEREND, D.; SAPIR, L. Monotonicity in Condorcet's Jury Theorem with dependent voters. **Social Choice and Welfare**, v. 28, p. 507-528, 2007.
- BLACK, D. On the Rationale of Group Decision-Making. **Journal of Political Economy**, v. 56, p. 23-34, 1948.
- BLACKORBY, C.; BOSSERT, W.; DONALDSON, D. Information Invariance in Variable-Population Social-Choice Problems. **International Economic Review**, v. 40, p. 403-422, 1999.
- BLACKORBY, C.; BOSSERT, W.; DONALDSON, D. **Population Issues in Social Choice Theory, Welfare Economics, and Ethics**. Cambridge: Cambridge University Press, 2005.
- BLAU, J. H. Liberal Values and Independence. **Review of Economic Studies**, v. 42, p. 395-401, 1975.
- BOLAND, P. J. Majority systems and the Condorcet jury theorem. **Statistician**, v. 38, p. 181-189, 1989.

- BOSSERT, W.; WEYMARK, J. A. Utility in social choice. *In*: BARBERÀ, P.; HAMMOND, J.; SEIDEL, C. (ed.). **Handbook of Utility Theory**. S. Boston: Kluwer, 1996. v. 2.
- BOSSERT, W.; QI, C. X.; WEYMARK, J. A. Extensive social choice and the measurement of group fitness in biological hierarchies. **Biology and Philosophy**, v. 28, p. 75-98, 2013.
- BOVENS, L.; RABINOWICZ, W. Democratic Answers to Complex Questions: An Epistemic Perspective. **Synthese**, v. 150, p. 131-153, 2006.
- BRAMS, S. J.; TAYLOR, A. D. **Fair Division: From Cake-Cutting to Dispute Resolution**. Cambridge: Cambridge University Press, 1996.
- BRANDT, F.; CONITZER, V.; ENDRISS, U. Computational Social Choice. **Multiagent Systems**. Cambridge: MIT Press, 2013, p. 213-283.
- BRIGHOUSE, H.; FLEURBAEY, M. Democracy and Proportionality. **Journal of Political Philosophy**, v. 18, p. 137-155, 2010.
- CHAPMAN, B. Rational Aggregation. **Politics, Philosophy and Economics**, v. 1, p. 337-354, 2002.
- CONDORCET, N. de. **Essay sur l'Application de l'Analyse à la Probabilité des Décisions Rendue à la Pluralité des Voix**. Paris, 1785,
- CONRADT, L.; LIST, C. Group decisions in humans and animals: a survey. **Philosophical Transactions of the Royal Society**, v. B 364, p. 719-742, 2009.
- CRAVEN, J. Liberalism and Individual Preferences. **Theory and Decision**, v. 14, p. 351-360, 1982.
- DESCHAMPS, R.; GEVERS, L. Leximin and utilitarian rules: A joint characterization. **Journal of Economic Theory**, v. 17, p. 143-163, 1978.
- DIETRICH, F. Judgment Aggregation: (Im)Possibility Theorems. **Journal of Economic Theory**, v. 126, p. 286-298, 2006.
- DIETRICH, F. A generalised model of judgment aggregation. **Social Choice and Welfare**, v. 28, p. 529-565, 2007.
- DIETRICH, F. The premises of Condorcet's jury theorem are not simultaneously justified. **Episteme**, v. 5, p. 56-73, 2008.
- DIETRICH, F.; LIST, C. A Model of Jury Decisions Where All Jurors Have the Same Evidence. **Synthese**, v. 142, p. 175-202, 2004.
- DIETRICH, F.; LIST, C. Judgment aggregation by quota rules: majority voting generalized. **Journal of Theoretical Politics**, v. 19, p. 391-424, 2007a.
- DIETRICH, F.; LIST, C. Arrow's theorem in judgment aggregation. **Social Choice and Welfare**, v. 29, p. 19-33, 2007b.

- DIETRICH, F.; LIST, C. Strategy-proof judgment aggregation. **Economics and Philosophy**, v. 23, p. 269-300, 2007c.
- DIETRICH, F.; LIST, C. **Opinion pooling on general agendas**, 2007d. Disponível em: <https://personal.lse.ac.uk/list/pdf-files/opinionpooling.pdf>. Acesso em: 13 jan. 2022.
- DIETRICH, F.; LIST, C. Judgment aggregation without full rationality. **Social Choice and Welfare**, v. 31, p. 15-39, 2008.
- DIETRICH, F.; LIST, C. Majority voting on restricted domains. **Journal of Economic Theory**, v. 145, p. 512-543, 2010a.
- DIETRICH, F.; LIST, C. The aggregation of propositional attitudes: towards a general theory. **Oxford Studies in Epistemology**, v. 3, p. 215-234, 2010b.
- DIETRICH, F.; MONGIN, P. The premise-based approach to judgment aggregation. **Journal of Economic Theory**, v. 145, p. 562-582, 2010.
- DIETRICH, F.; SPIEKERMANN, K. Epistemic Democracy with Defensible Premises. **Economics and Philosophy**, v. 29, n.1, p. 87-120, 2013.
- DOKOW, E.; HOLZMAN, R. Aggregation of binary evaluations. **Journal of Economic Theory**, v. 145, p. 495-511, 2010a.
- DOKOW, E.; HOLZMAN, R. Aggregation of binary evaluations with abstentions. **Journal of Economic Theory**, v. 145, p. 544-561, 2010b.
- DOKOW, E.; HOLZMAN, R. Aggregation of non-binary evaluations. **Advances in Applied Mathematics**, v. 45, p. 487-504, 2010c.
- DOWDING, K.; VAN HEES, M. The Construction of Rights. **American Political Science Review**, v. 97, p. 281-293, 2003.
- DOWDING, K.; VAN HEES, M. In Praise of Manipulation. **British Journal of Political Science**, v. 38, p. 1-15, 2008.
- DRYZEK, J.; LIST, C. Social Choice Theory and Deliberative Democracy: A Reconciliation. **British Journal of Political Science**, v. 33, p. 1-28, 2003.
- ELSHOLTZ, C.; LIST, C. A Simple Proof of Sen's Possibility Theorem on Majority Decisions. **Elemente der Mathematik**, v. 60, p. 45-56, 2005.
- ELSTER, J. Excessive Ambitions (II). **Capitalism and Society**, v. 8, n. 1, art. 1, 2013.
- ESTLUND, D. Opinion Leaders, Independence, and Condorcet's Jury Theorem. **Theory and Decision**, v. 36, p. 131-162, 1994.
- FEDDERSEN, T. J.; PESENDORFER, W. Convicting the Innocent. **American Political Science Review**, v. 92, p. 23-35, 1998.
- GAERTNER, W. **Domain Conditions in Social Choice Theory**. Cambridge: Cambridge University Press, 2001.

- GAERTNER, W. De jure naturae et gentium: Samuel von Pufendorf's contribution to social choice theory and economics. **Social Choice and Welfare**, v. 25, p. 231-241, 2005.
- GAERTNER, W.; PATTANAIK, P. K.; SUZUMURA, K. Individual Rights Revisited. **Economica**, v. 59, p. 161-177, 1992.
- GAERTNER, W.; SCHOKKAERT, E. **Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice**. Cambridge: Cambridge University Press, 2012.
- GÄRDENFORS, P. An Arrow-like theorem for voting with logical consequences. **Economics and Philosophy**, v. 22, p. 181-190, 2006.
- GEHRLEIN, W. V. Condorcet's Paradox. **Theory and Decision**, v. 15, p. 161-197, 1983.
- GENEST, C.; ZIDEK, J. V. Combining Probability Distributions: A Critique and Annotated Bibliography. **Statistical Science**, v. 1, p. 113-135, 1986.
- GIBBARD, A. **Social Choice and the Arrow Conditions**. 1969. Disponível em: <http://www-personal.umich.edu/~gibbard/Gibbard1968-Social-Choice-Arrow-Conditions.pdf>. Acesso em: 13 jan. 2022.
- GIBBARD, A. Manipulation of voting schemes: a general result. **Econometrica**, v. 41, p. 587-601, 1973.
- GIGLIOTTI, G. A. Comment on Craven. **Theory and Decision**, v. 21, p. 89-95, 1986.
- GILBOA, I.; SAMET, D.; SCHMEIDLER, D. Utilitarian Aggregation of Beliefs and Tastes. **Journal of Political Economy**, v. 112, p. 932-938, 2004.
- GOLDMAN, A. Group Knowledge versus Group Rationality: Two Approaches to Social Epistemology. *Episteme*. **A Journal of Social Epistemology**, v. 1, p. 11-22, 2004.
- GOLDMAN, A. Why Social Epistemology Is Real Epistemology. **Social Epistemology**. Oxford: Oxford University Press, 2010.
- GOODIN, R. E.; LIST, C. A Conditional Defense of Plurality Rule: Generalizing May's Theorem in a Restricted Informational Environment. **American Journal of Political Science**, v. 50, p. 940-949, 2006.
- GROFMAN, B.; OWEN, G.; FELD, S. L. Thirteen theorems in search of the truth. **Theory and Decision**, v. 15, p. 261-278, 1983.
- GUILBAUD, G. T. Theories of the General Interest, and the Logical Problem of Aggregation. **Readings in Mathematical Social Science**. Cambridge: MIT Press, [1952] 1966, p. 262-307.
- HARBOUR, D.; LIST, C. Optimality Theory and the problem of constraint aggregation. In: BHATT, R.; HAWLEY, P.; HACKL, M.; MAITRA, I. (ed.). **MIT Working**

- Papers in Linguistics and Philosophy 1:** The Linguistics/Philosophy Interface. Cambridge: MITWPL, 2000, p. 175-213.
- HARRISON, G. W.; MCDANIEL, T. Voting Games and Computational Complexity. **Oxford Economic Papers**, v. 60, p. 546-565, 2008.
- HAUSMAN, D. The Impossibility of Interpersonal Utility Comparisons. **Mind**, v. 104, p. 473-490, 1995.
- HURLEY, S. Supervenience and the Possibility of Coherence. **Mind**, v. 94, p. 501-525, 1985.
- INADA, K.-I. A Note on the Simple Majority Decision Rule. **Econometrica**, v. 32, p. 525-531, 1964.
- KANAZAWA, S. A brief note on a further refinement of the Condorcet Jury Theorem for heterogeneous groups. **Mathematical Social Sciences**, v. 35, p. 69-73, 1998.
- KNIGHT, J.; JOHNSON, J. Aggregation and Deliberation: On the Possibility of Democratic Legitimacy. **Political Theory**, v. 22, p. 277-296, 1994.
- KONIECZNY, S.; PINO PÉREZ, R. Merging Information Under Constraints: A Logical Framework. **Journal of Logic and Computation**, v. 12, p. 773-808, 2002.
- KORNHAUSER, L. A.; SAGER, L. G. Unpacking the Court. **Yale Law Journal**, v. 96, p. 82-117, 1986.
- KROEDEL, T.; HUBER, F. Counterfactual Dependence and Arrow. **Noûs**, v. 47, n. 3, p. 453-466, 2013.
- LADHA, K. The Condorcet Jury Theorem, Free Speech and Correlated Votes. **American Journal of Political Science**, v. 36, p. 617-634, 1992.
- LEHRER, K.; WAGNER, C. **Rational Consensus in Science and Society**. Dordrecht/Boston: Reidel, 1981.
- LIST, C. A Note on Introducing a 'Zero-Line' of Welfare as an Escape-Route from Arrow's Theorem. **Pacific Economic Review**, n. 6, p. 223-238, 2001.
- LIST, C. **The epistemology of special majority voting**. Working paper, London School of Economics. 2003a. Disponível em: <https://personal.lse.ac.uk/list/PDF-files/proportion3107.pdf>. Acesso em: 13 jan. 2022.
- LIST, C. Are Interpersonal Comparisons of Utility Indeterminate? **Erkenntnis**, v. 58, p. 229-260, 2003b.
- LIST, C. A Possibility Theorem on Aggregation over Multiple Interconnected Propositions. **Mathematical Social Sciences**, v. 45, p. 1-13, 2003c. (with Corrigendum in **Mathematical Social Sciences**, v. 52, p. 109-110).
- LIST, C. Multidimensional Welfare Aggregation. **Public Choice**, v. 119, p. 119-142, 2004a.

- LIST, C. A Model of Path-Dependence in Decisions over Multiple Propositions. **American Political Science Review**, v. 98, p. 495-513, 2004b.
- LIST, C. The Discursive Dilemma and Public Reason. **Ethics**, v. 116, p. 362-402, 2006.
- LIST, C. The Logical Space of Democracy. **Philosophy and Public Affairs**, v. 39, p. 262-297, 2011.
- LIST, C. The Theory of Judgment Aggregation: An Introductory Review. **Synthese**, v. 187, p. 179-207, 2012.
- LIST, C.; GOODIN, R. E. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. **Journal of Political Philosophy**, v. 9, p. 277-306, 2001.
- LIST, C.; LUSKIN, R. C.; FISHKIN, J. S.; MCLEAN, I. Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. **Journal of Politics**, v. 75, p. 80-95, 2013.
- LIST, C.; PETTIT, P. Aggregating Sets of Judgments: An Impossibility Result. **Economics and Philosophy**, v. 18, n. 1, p. 89-110, 2002.
- LIST, C.; PETTIT, P. Aggregating Sets of Judgments: Two Impossibility Results Compared. **Synthese**, v. 140, p. 207-235, 2004.
- MAY, K. O. A set of independent, necessary and sufficient conditions for simple majority decision. **Econometrica**, v. 20, p. 680-684, 1952.
- MAY, K. O. Intransitivity, Utility, and the Aggregation of Preference Patterns. **Econometrica**, v. 22, p. 13, 1954.
- MCCONWAY, K. J. Marginalization and Linear Opinion Pools. **Journal of the American Statistical Association**, v. 76, p. 410-414, 1981.
- MCLEAN, I. The Borda and Condorcet principles: Three medieval applications. **Social Choice and Welfare**, v. 7, p. 99-108, 1990.
- MCLEAN, I.; HEWITT, F. (ed.). **Condorcet: Foundations of Social Choice and Political Theory**. Cheltenham: Edward Elgar Publishing, 1994.
- MCLEAN, I. S.; MCMILLAN, A.; MONROE, B. L. Duncan Black and Lewis Carroll. **Journal of Theoretical Politics**, v. 7, p. 107-123, 1995.
- MCLEAN, I. S.; MCMILLAN, A.; MONROE, B. L. (ed.). **A Mathematical Approach to Proportional Representation: Duncan Black on Lewis Carroll**. Dordrecht: Kluwer, 1996.
- MCLEAN, I.; URKEN, A. B. (ed.). **Classics of Social Choice**. Ann Arbor: University of Michigan Press, 1995.
- MILLER, D. Deliberative Democracy and Social Choice. **Political Studies**, v. 40, special issue, p. 54-67, 1992.

- MILLER, M. K.; OSHERSON, D. Methods for distance-based judgment aggregation. **Social Choice and Welfare**, v. 32, p. 575-601, 2009.
- MONJARDET, B. Social choice theory and the Centre de Mathématique Sociale: some historical notes. **Social Choice and Welfare**, v. 25, p. 433-456, 2005.
- MONGIN, P. Consistent Bayesian aggregation. **Journal of Economic Theory**, v. 66, p. 313-351, 1995.
- MONGIN, P. **Spurious Unanimity and the Pareto Principle**. Paper presented at the Conference on Utilitarianism, New Orleans, March, 1997. Disponível em: [http://eprints.lse.ac.uk/23942/1/Spurious_Unanimity_and_the_Pareto_Principle_\(LSERO\).pdf](http://eprints.lse.ac.uk/23942/1/Spurious_Unanimity_and_the_Pareto_Principle_(LSERO).pdf). Acesso em: 14 jan. 2022.
- MONGIN, P. Factoring Out the Impossibility of Logical Aggregation. **Journal of Economic Theory**, v. 141, p. 100-113, 2008.
- MORREAU, M. It simply does not add up: Trouble with overall similarity. **Journal of Philosophy**, v. 107, p. 469-490, 2010.
- MORREAU, M. Theory Choice and Social Choice: Kuhn Vindicated. **Mind**, v. 124, n. 493, p. 239-262, 2015.
- MOULIN, H. On Strategy-Proofness and Single Peakedness. **Public Choice**, v. 35, p. 437-455, 1980.
- MOULIN, H. **Fair Division and Collective Welfare**. Cambridge: MIT Press, 2004.
- MUELLER, D. C. **Public Choice III**. Cambridge: Cambridge University Press, 2003.
- NEHRING, K. Arrow's theorem as a corollary. **Economics Letters**, v. 80, p. 379-382, 2003.
- NEHRING, K.; PUPPE, C. **Strategyproof Social Choice on Single-Peaked Domains**: Possibility, Impossibility and the Space Between. California: University of California at Davis, 2002.
- NEHRING, K.; PUPPE, C. The structure of strategy-proof social choice - Part I: General characterization and possibility results on median spaces. **Journal of Economic Theory**, v. 135, p. 269-305, 2007.
- NEHRING, K.; PUPPE, C. Abstract Arrowian Aggregation. **Journal of Economic Theory**, v. 145, p. 467-494, 2010.
- OKASHA, S. Individuals, groups, fitness and utility: multi-level selection meets social choice theory. **Biology and Philosophy**, v. 24, p. 561-584, 2009.
- OKASHA, S. Theory Choice and Social Choice: Kuhn versus Arrow. **Mind**, v. 120, p. 83-115, 2011.
- OOGHE, E.; LAUWERS, L. Non-dictatorial extensive social choice. **Economic Theory**, v. 25, p. 721-743, 2005.

- PARFIT, D. **Reasons and Persons**. Oxford: Oxford University Press, 1984.
- PAULY, M.; VAN HEES, M. Logical Constraints on Judgment Aggregation. **Journal of Philosophical Logic**, v. 35, p. 569-585, 2006.
- PEARL, J. **Causality: models, reasoning, and inference**. Cambridge: Cambridge University Press, 2000.
- PETTIT, P. Deliberative Democracy and the Discursive Dilemma. **Philosophical Issues**, v. 11, p. 268-299, 2001.
- PIGOZZI, G. Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. **Synthese**, v. 152, p. 285-298, 2006.
- POISSON, S. D. **Recherches sur la probabilité des jugements en matière criminelle et en matière civile: précédées des règles générales du calcul des probabilités**, 1837.
- RAWLS, J. **A Theory of Justice**. Cambridge: Harvard University Press, 1971.
- RAWLS, J. **Uma Teoria da Justiça**. São Paulo: Martins Fontes, 2016.
- REGENWETTER, M.; B. GROFMAN, B.; MARLEY, A. A. J.; TSETLIN, I. **Behavioral Social Choice**. Cambridge: Cambridge University Press, 2006.
- RIKER, W. **Liberalism against Populism**. San Francisco: W.H. Freeman and Co., 1982.
- ROBERTS, K. Valued Opinions or Opinionated Values: The Double Aggregation Problem. **Choice, Welfare and Development: A Festschrift in Honour of Amartya Sen**. Oxford: Oxford University Press, 1995, p. 141-165.
- SAARI, D. G. The Borda dictionary. **Social Choice and Welfare**, v. 7, p. 279-317, 1990.
- SALLES, M. The history of Social Choice. **Social Choice and Welfare**, v. 25, p. 229-564, 2005.
- SATTERTHWAITE, M. Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. **Journal of Economic Theory**, v. 10, p. 187-217, 1975.
- SEN, A. K. A Possibility Theorem on Majority Decisions. **Econometrica**, v. 34, p. 491-499, 1966.
- SEN, A. K. The Impossibility of a Paretian Liberal. **Journal of Political Economy**, v. 78, p. 152-157, 1970a.
- SEN, A. K. **Collective Choice and Social Welfare**. San Francisco: Holden-Day, 1970b.
- SEN, A. K. **Escolha Coletiva e Bem-estar Social**. São Paulo: Almedina, 2019.
- SEN, A. K. On weights and measures: informational constraints in social welfare analysis. **Econometrica**, v. 45, p. 1539-1572, 1977.
- SEN, A. K. **Choice, Welfare and Measurement**. Oxford: Blackwell, 1982.
- SEN, A. K. Liberty and social choice. **Journal of Philosophy**, v. 80, p. 5-28, 1983.

- SEN, A. K. **The Possibility of Social Choice**. Nobel lecture, December 8, 1998, Stockholm. Disponível em: <https://www.nobelprize.org/prizes/economic-sciences/1998/sen/lecture/>. Acesso em: 14 jan. 2022.
- SHAPLEY, L.; GROFMAN, B. Optimizing group judgment accuracy in the presence of interdependencies. **Public Choice**, v. 43, p. 329-343, 1984.
- STEGENGA, J. An impossibility theorem for amalgamating evidence. **Synthese**, v. 190, n. 2, p. 2391-2411, 2013.
- SUPPES, P. The pre-history of Kenneth Arrow's social choice and individual values. **Social Choice and Welfare**, v. 25, p. 319-226, 2005.
- THOMSON, W. On the axiomatic method and its recent applications to game theory and resource allocation. **Social Choice and Welfare**, v. 18, p. 327-386, 2000.
- TSETLIN, I.; REGENWETTER, M.; GROFMAN, B. The Impartial Culture Maximizes the Probability of Majority Cycles. **Social Choice and Welfare**, v. 21, p. 387-398, 2003.
- VACCA, R. Opinioni Individuali e Deliberazioni Collettive. **Rivista Internazionale di Filosofia del Diritto**, p. 52-59, 1921.
- WARD, B. Majority Voting and Alternative Forms of Public Enterprises. **The Public Economy of Urban Communities**. Baltimore: Johns Hopkins Press, 1965.
- WEYMARK, J. The Normative Approach to the Measurement of Multidimensional Inequality. **Inequality and Economic Integration**. London: Routledge, 2006, p. 303-328.
- WILSON, R. On the Theory of Aggregation. **Journal of Economic Theory**, v. 10, p. 89-99, 1975.

Ética dos Negócios*

Jeffrey Moriarty (Universidade de Bentley)**

Tradução: Bruno Aislã Gonçalves dos Santos (UNICENTRO)

Revisão: Ramiro de Ávila Peres (BCB)

Um negócio (ou empresa) é uma organização produtiva – uma organização cujo propósito é criar bens e serviços para venda, normalmente para o lucro. Negócios são também uma atividade. Uma entidade (por exemplo, uma pessoa,

* MORIARTY, J. Business Ethics. In: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Fall Edition. Stanford, CA: The Metaphysics Research Lab, 2017. Disponível em: <https://plato.stanford.edu/archives/fall2017/entries/ethics-business/>. Acesso em: 12 jan. 2022.

The following is the translation of the entry on “Business Ethics” by Jeffrey Moriarty in the Stanford Encyclopedia of Philosophy. The translation follows the version of the entry in the SEP’s archives at <<https://plato.stanford.edu/archives/fall2017/entries/ethics-business/>>. This translated version may differ from the current version of the entry, which may have been updated since the time of this translation. The current version is located at a <https://plato.stanford.edu/entries/ethics-business/> We’d like to thank the Editors of the Stanford Encyclopedia of Philosophy, mainly Prof. Dr. Edward Zalta, for granting permission to translate and to publish this entry.

** Pelas sugestões muito úteis em relação a esse verbete, e sua versão anterior, agradeço Dorothea Baur, George Brenkert, Jason Brennan, Matt Caulfield, David Dick, Anca Gheaus, Keith Hankins, Edwin Hartman, Laura Hartman, Lisa Herzog, David Jacobs, Woon Hyuk Jay Jang, Peter Jaworski, Xavier Landes, Chris MacDonald, Emilio Marti, Dominic Martin, Pierre-Yves Néron, Eric Orts, Katinka Quintelier, Sareh Pouryousefi, Amy Sepinwall, Kenneth Silver, Abraham Singer, Alejo José G. Sison, Cindy Stark, Chris Surprenant, Kevin Vallier e Hasko von Kriegstein.

uma organização) “faz negócios” com outra quando elas trocam um bem ou serviço com caráter oneroso. Ética dos negócios pode então ser entendida como o estudo das dimensões éticas das organizações produtivas e das atividades empresariais/comerciais. Isso inclui a análise ética da produção, distribuição, *marketing*, venda e consumo de bem e serviços (Ver também DONALDSON; WALSH, 2015).

As questões de ética dos negócios são importantes e relevantes para todos. Isso se dá porque quase todos nós “fazemos negócios” – por exemplo, nos engajamos em transações comerciais – quase todos os dias. Além disso, muitos de nós gastam a maior parte de nossas vidas nos engajando, ou nos preparando para nos engajar, em atividades produtivas, por conta própria ou como parte de uma organização produtiva. Atividades comerciais moldam o mundo em que vivemos, algumas vezes para o bem, outras vezes para o mal.

A ética dos negócios é um amplo campo de pesquisa. Filósofos desde Aristóteles até Rawls têm defendido posições acerca de tópicos que podem ser entendidos como partes da ética dos negócios. No momento, há pelo menos cinco revistas especializadas e devotadas a esse campo (*Business Ethics Quarterly*, *Business Ethics: A European Review*, *Business & Society*, *Business & Society Review*, *Journal of Business Ethics*), novos trabalhos aparecem na filosofia de forma geral e, também, em revistas especializadas de ciências sociais.

Este texto sumariza importantes pesquisas sobre questões centrais da ética dos negócios, incluindo: No interesse de quem as empresas deveriam ser administradas? Quem deveria administrá-las? O que as empresas devem aos seus funcionários, e o que os funcionários devem para suas empresas? Quais regras morais deveriam guiar o engajamento das empresas junto aos consumidores? As empresas deveriam tentar resolver problemas sociais? Quais as responsabilidades que elas têm em relação aos seus fornecedores? Qual o papel que as empresas deveriam ter no processo político? Dada a vastidão do campo, por necessidade, certas questões de ética nos negócios não são abordadas aqui.

1. Variedades da ética dos negócios

Muitas pessoas envolvidas nas atividades de negócios, incluindo contadores e advogados, são profissionais. Como tal, eles estão ligados por códigos de conduta promulgadas por sociedades profissionais. Muitas empresas também têm códigos de condutas detalhados desenvolvidos e reforçados por equipes de ética e de

conformidade. A ética dos negócios pode ser entendida como o estudo das práticas profissionais, ou seja, como o estudo do conteúdo, desenvolvimento, gerenciamento e efetividade dos códigos de conduta criados para guiar as ações das pessoas engajadas em atividade comerciais. Este texto não considera esta forma de ética dos negócios. Ao invés disso, ele considera a ética dos negócios como uma disciplina acadêmica.

A ética dos negócios como disciplina acadêmica é formada por cientistas sociais e teóricos normativos. Isso é refletido nos participantes de conferências acadêmicas sobre ética nos negócios e nos tipos de artigos publicados em revistas especializadas de ética dos negócios. Cientistas sociais – aqueles, que neste ponto, compreendem o maior grupo dentro deste campo – focam o estudo da ética dos negócios de modo descritivo. Eles tentam responder questões como: O desempenho social corporativo melhora o desempenho financeiro, ou seja, a ética compensa financeiramente (MARGOLIS; WALSH, 2003; ORLITZKY; SCHMIDT; RYNES, 2003)? E o que faz as pessoas engajarem-se em comportamentos antiéticos (BAZERMAN; TENBRUNSEL, 2011; WERHANE *et al.*, 2013)? Eu não considerarei tais questões aqui. Este texto foca questões normativas da ética dos negócios, sendo a maioria uma variação da questão: O que é ético e antiético de se fazer nos negócios?

Considerada somente como um empreendimento normativo, a ética dos negócios é – assim como muitas áreas da ética aplicada – extraída de uma variedade de disciplinas, incluindo ética, filosofia política, economia, psicologia, direito e políticas públicas. Isso se dá porque as soluções para comportamentos antiéticos nos negócios podem tomar várias formas, desde recomendações dirigidas a indivíduos particulares para que mudem seu comportamento até novas leis, políticas e regulamentações. Fazer ética dos negócios bem significa estar familiarizado com os resultados nestas disciplinas ou, pelo menos, estar ciente das lacunas existentes em seu próprio conhecimento.

Estudiosos normativistas da ética dos negócios (doravante, o qualificador “normativo” será assumido) costumam fazer certas suposições acerca do linguajar econômico. Uma é que os meios de produção podem ser de propriedade privada. Uma segunda suposição é que mercados – apresentando trocas voluntárias entre compradores e vendedores a preços determinados mutuamente – devem desempenhar um papel importante na alocação de recursos. Aqueles que negam essas suposições irão ver alguns debates em ética dos negócios (e.g., sobre a propriedade e o controle de uma empresa ou sobre propaganda) como equivocados.

Algumas organizações “fazem negócios” – no sentido de trocar um bem ou serviço com caráter oneroso – com o objetivo de buscar o lucro, e algumas não

o fazem. Merck e Wal-Mart são exemplos do primeiro tipo de organizações. A Universidade de Princeton e o Museu Metropolitano de Arte são exemplos do segundo tipo. Eticistas dos negócios estão preocupados algumas vezes com as atividades das organizações sem fins lucrativos, mas, mais comumente, focam organizações com fins lucrativos. De fato, a maioria das pessoas provavelmente entende as empresas como organizações que visam ao lucro.

2. Agência Moral Corporativa

Uma maneira de pensar sobre a ética nos negócios é pensar em termos de obrigações morais dos agentes engajados em atividade comerciais. Quem é o agente moral? Essa questão é tipicamente descrita como uma questão de “agência moral corporativa” ou de “responsabilidade moral corporativa”. Aqui “corporativo” não refere a corporação como uma entidade legal, mas como um coletivo ou grupo de indivíduos. Para ser preciso, a questão é se empresas são agentes morais, considerados moralmente responsáveis como (**enquanto**) empresas, não considerados enquanto agregados de membros individuais da empresa.

Na literatura da ética dos negócios, French é um pensador seminal sobre esse tópico. Em seus primeiros trabalhos (1979, 1984), ele argumenta que empresas são moralmente responsáveis pelo que fazem, e, portanto, deve ser vista como pessoas morais “de pleno direito”. Ele baseia essa conclusão sobre afirmações de que empresas tem estruturas internas de tomada de decisão, através das quais elas (1) causam eventos e; (2) agem intencionalmente. Algumas respostas aos primeiros trabalhos de French aceitaram a afirmação de que empresas são agentes morais, mas negam que empresas são pessoas morais. Donaldson (1982) defende que empresas não podem ser pessoas porque falta a elas importantes capacidades humanas, tais como, a habilidade de perseguir sua própria felicidade (Ver também WEHANE, 1985). Outros respondem negando que empresas sejam agente morais (também). Em seu último trabalho, French (1995) retrocedeu em sua afirmação de que empresas são pessoas morais, embora não tenha feito o mesmo acerca da sua afirmação de que elas são agentes morais.

As discussões da agência moral corporativa e da responsabilidade moral tem desaparecido da literatura da ética dos negócios (como em 2016). Mas elas continuam a receber atenção na literatura filosófica geral, onde elas são tratadas com um alto grau de sofisticação. Na filosofia de forma geral, o foco está sobre

coletivos de forma mais geral, com as empresas comerciais desempenhando o papel com exemplos de coletividade. Como na literatura de ética dos negócios, na literatura filosófica geral, a questão central é: Quais são as condições para a agência e responsabilidade moral, tais que as coletividades **enquanto** coletividades, incluindo empresas, as satisfazem ou não? Um número de escritores, incluindo Copp (2006), Hess (2014), e List e Pettit (2011), acreditam que empresas podem ser agente morais. Essa visão tem um forte apelo intuitivo. Nós rotineiramente dizemos coisas como: “Costco trata seus empregados bem” ou “BP poluiu o Golfo do México”, e ao fazer isso nós parecemos concordar com a agência e responsabilidade moral das empresas tomadas nelas mesmas. Do outro lado estão os escritores que negam que empresas podem ser agentes morais, tais como Gilbert (1989), S. Miller (2006), e Rønnegard (2015). Uma afirmação apresentada desse lado é que a agência exige intenção, e as empresas não são o tipo de coisa que pode ter intenções (MILLER, 2006). A maneira comum de falarmos acerca de agência e responsabilidade das empresas pode ser metafórica, ou um atalho para referir a agência e responsabilidade dos indivíduos dentro das empresas (para discussões desses pontos, confira os verbetes da Stanford **collective responsibility**, **collective intentionality**, e **shared agency**).

Enquanto a questão de se as próprias empresas são agentes morais é de interesse teórico, sua importância prática é incerta. Talvez a própria BP haja sido moralmente responsável por poluir o Golfo do México. Talvez certos indivíduos que trabalham na BP o sejam. O que está por trás disso? De acordo com Hasnas (2012), muito pouco. Empresas como BP podem ser legalmente obrigadas a pagarem restituições por danos que elas casam mesmo se elas não são moralmente responsáveis por eles. O que atribuir agência e responsabilidade às empresas nos permite fazer, segundo Hasnas, é culpá-las e puni-las. Mas, ele argumenta, nós não deveríamos nos engajar nessa prática. Phillips (1995), em contraste, argumenta que em alguns casos nenhum empregado individual em uma empresa é responsável por um dano que uma empresa causa. Na medida em que faz sentido – e geralmente faz, ele acredita – atribuir responsabilidade pelo dano, ele deve ser atribuído à própria empresa. Na visão de Phillips, agência moral corporativa torna possível culpar o comportamento danoso onde ele, de outro modo, não o seria. Como a reputação corporativa pode ser um ativo ou passivo significativo (ROBERTS; DOWLING, 2002), isso oferece um incentivo para a empresa exercer o devido cuidado em suas operações (Ver também RÖNNEGARD, 2015).

3. Os fins e os meios da governança corporativa.

Há um significativo debate acerca dos fins e meios da governança corporativa, ou seja, sobre para quem as empresas deveriam ser gerenciadas e, quem deveria (em última instância) gerenciá-las. Grande parte desse debate é realizado com as grandes empresas de capital aberto.

3.1 Os fins: Primazia dos acionistas ou equilíbrio entre as partes interessadas?⁴⁰

Há duas visões principais sobre os fins adequados da governança corporativa. De acordo com uma visão, empresas deveriam ser gerenciadas nos melhores interesses dos acionistas. É tipicamente assumido que o gerenciamento das empresas nos melhores interesses dos acionistas requer maximizar a riqueza deles. Essa visão é frequentemente chamada de “primazia dos acionistas” (STOUT, 2002) ou – a fim de contrastá-las mais diretamente com sua principal rival (que será discutida abaixo) “teoria do acionista”. De maneira confusa, o rótulo “primazia do acionista” às vezes é usado – por exemplo, por Bainbridge [2008] – para se referir à visão de que os acionistas devem ter o controle final sobre a empresa.

Alguns poucos escritores argumentam pela primazia dos acionistas a partir de fundamentos deontológicos. Nessa argumentação, acionistas são donos da empresa, e contratam administradores para gerenciá-las para eles sob a condição de que a empresa seja gerida em seus interesses. A primazia dos acionistas é então baseada em uma promessa que administradores fazem aos acionistas (FRIEDMAN, 1970; HASNAS, 1998). Em resposta, alguns argumentam que os acionistas não

⁴⁰ N.T.: No original em inglês os termos que se referem a “acionistas” e a “as partes interessadas” são, respectivamente, “shareholders” e “stakeholders”. É comum encontrar na bibliografia especializada em português os termos em inglês, ao invés de uma tradução. Aqui opto pela tradução, pois este é um texto introdutório, porém, há certa discordância sobre a tradução dos termos, principalmente, de “stakeholders”. Opto por traduzir “stakeholders” por “partes interessadas”, pois, como o leitor perceberá no decorrer do texto (principalmente em 3), há discussão sobre quem são as partes interessadas quando se trata da governança corporativa. Além disso, esse termo é utilizado em algumas normas regulatórias – sobre, por exemplo, responsabilidade socioambiental de instituições financeiras.

são donos das empresas. Eles detêm as ações, um tipo de título de dívida corporativa (BAINBRIDGE, 2008; STOUT, 2002); a empresa em si não pode ser propriedade (STRUDLER, 2017). Outros argumentam que os administradores não fazem, explícita ou implicitamente, qualquer promessa aos acionistas para gerenciar a empresa de certa maneira (BOATRIGT, 1994). Muitos escritores argumentam pela primazia dos acionistas sob fundamentos consequencialistas. Sob esse argumento, gerenciar empresas nos interesses dos acionistas é mais eficiente que as gerenciar de outra maneira (HANSMANN; KRAAKAMAN, 2001; JESSEN, 2002). Para sustentar isso, alguns argumentam que, se os administradores não tiverem um objetivo único que seja claro e mensurável – viz., maximizar o valor para as partes interessadas –, eles terão uma oportunidade aprimorada de negociação própria (STOUT, 2002). Argumentos consequencialistas pela primazia dos acionistas encontram problemas que afligem muitas versões do consequencialismo: ao exigir que todas as empresas sejam gerenciadas de uma certa maneira, isso não permite espaço suficiente para a escolha pessoal (HUSSAIN, 2012). A maioria pensa que as pessoas deveriam ser capazes de perseguir seus projetos, incluindo projetos econômicos, que lhes são relevantes, mesmo que esses projetos não maximizem o bem-estar.

A segunda visão principal sobre os fins apropriados da governança corporativa é dada pela teoria das partes interessadas. Essa teoria foi a primeira oferecida por Freeman na década de 80 (FREEMAN, 1984; FREEMAN; REED, 1983) e então refinadas por ele e vários colaboradores nos próximos 30 anos (por exemplo, e.g. FREEMAN *et. al.*, 2010; JONES; WICKS; FREEMAN, 2002). De acordo com a teoria das partes interessadas – ou, pelo menos, as primeiras formulações da teoria – ao invés de administrar somente uma empresa nos melhores interesses dos acionistas, administradores deveriam buscar “balancear” os interesses de todas as partes interessadas, onde um interessado é qualquer um que tem uma “participação” ou interesse (incluindo interesse financeiro) na empresa.

Para seus críticos, a teoria das partes interessadas pareceu ter sido insuficientemente articulada e fracamente defendida. A respeito da articulação, uma questão que tem sido pressionada é: Quem são as partes interessadas (ORTS; STRUDLER, 2002, 2009)? Os grupos mais comumente identificados são os acionistas, empregados, a comunidade, os fornecedores e consumidores. Mas outros grupos têm participações na empresa, incluindo credores, o governo e competidores. Faz uma grande diferença onde a linha é traçada, mas teóricos das partes interessadas não têm oferecido uma razão clara para traçar uma linha em um ponto e não em outro. Outra questão é: o que significa “balancear” os interesses de todas as partes

interessadas – diferente de nem sempre dar precedência a interesses dos acionistas (ORTS; STRUDLER, 2009)? No que diz respeito à defesa, os críticos se perguntam qual é a justificativa para administrar as empresas no interesse de todas as partes interessadas. Num ponto, Freeman (1984) oferece um argumento instrumental para sua visão, afirmando que balancear os interesses das partes interessadas é melhor estrategicamente para empresa que maximizar a riqueza dos acionistas. (Isso é precisamente o que defensores da primazia dos acionistas dizem acerca dessa visão). Num outro ponto, ele dá um argumento que apela para a justiça como equidade de Rawls (EVAN; FREEMAN, 1988; CHILD; MARCOUX, 1999).

Nos anos recentes, têm sido levantadas questões sobre se a teoria das partes interessadas é apropriadamente vista como uma opositora genuína à primazia dos acionistas, ou se é mesmo apropriado chamá-la de uma “teoria”. Em um artigo, Freeman e colaboradores dizem que a teoria das partes interessadas é simplesmente “o corpo de pesquisa [...] no qual a ideia de ‘interessados’ desempenha um papel crucial” (JONES *et al.*, 2002). Em outro, Freeman descreve a teoria das partes interessadas como “um gênero de histórias sobre como nós poderíamos viver” (1994, p. 413). Pode ser, como Norman (2013) diz, que a parte interessada é agora mais considerada mais uma “mentalidade”, isto é, uma maneira de olhar para a empresa que enfatiza sua incorporação em uma rede de relacionamentos.

É importante perceber que a resolução do debate entre teóricos dos acionistas e das partes interessadas (como concebemos os últimos) não resolverá todas ou mesmo a maioria das questões éticas nos negócios. Isso se dá porque é um debate sobre os fins da governança corporativa; isto não pode responder todas as questões sobre as **restrições morais** que devem ser observadas na busca desses fins (GOODPASTER, 1991; NORMAN, 2013). Nem a primazia dos acionistas nem a teoria das partes interessadas são plausivelmente interpretadas como a visão de que administradores corporativos deveriam fazer o **possível** para maximizar a riqueza dos acionistas e balancear todos os interesses das partes interessadas, respectivamente. Ao invés disso, essas visões deveriam ser interpretadas como visões do que os administradores deveriam fazer seja o que for **moralmente permissível** para realizar esses fins. Uma grande parte da ética dos negócios é tentar determinar o que é moralmente permissível dentro desse domínio.

3.2 Meios: controle pelos acionistas ou também por outros?

Respostas para a questão sobre os meios da governança corporativa frequentemente espelham respostas para a questão sobre os fins da governança corporativa. Frequentemente, a melhor maneira de assegurar que uma empresa seja administrada nos interesses de uma certa parte P é dar a P controle sobre ela. Inversamente, justificações do porquê a empresa deveria ser administrada no interesse de P algumas vezes apelam para o direito de P de controlá-la.

Assim Friedman (1970) pensa que a propriedade da firma pelos acionistas lhes dá o direito de controlá-la (os quais podem usá-la para assegurar que seja administrada em seus interesses). Podemos ver os direitos de controle dos acionistas como se seguindo analiticamente do conceito de propriedade. Deter uma coisa é ter um pacote de direitos em relação a essa coisa. Um dos “incidentes” padrão de propriedade é o controle.

Como dito, em anos recentes a ideia de que a empresa é algo que pode ser possuído tem sido desafiada (BAINBRIDGE, 2008; STRUDLER, 2017). Mas argumentos contratualistas para o controle dos acionistas da empresa têm sido construídos, os quais não dependem da suposição de propriedade da empresa. Tudo o que é suposto nesses argumentos é que alguma pessoa detém capital, e os outros detém mão-de-obra. O capital pode “contratar” mão-de-obra (e outros insumos de produção), nos termos que elabora, ou a mão-de-obra pode “contratar” o capital, nos seus próprios termos, com a sociedade estabelecendo limites quanto aos termos. Acontece que, na maioria dos casos, o capital contrata mão-de-obra. Esses pontos são enfatizados especialmente por aqueles que consideram a empresa como um “nexo de contratos” entre várias partes (EASTERBROOK; FISCHER, 1996; JENSEN; MECKLING, 1976).

Muitos escritores acham esse resultado preocupante. Mesmo que a estrutura administrativa na maioria das empresas seja de algum modo acordada, eles dizem que é injusta de outras maneiras. Anderson (2015) caracteriza regimes tradicionais de governança corporativa como opressivos e ditaduras privadas opressivas e irresponsáveis. Para chamar a atenção para essa injustiça, esses escritores pedem várias formas de participação dos trabalhadores na tomada de decisões administrativas, incluindo a capacidade dos trabalhadores de rejeitar diretrizes arbitrárias dos administradores (HSIEH, 2005), a co-determinação entre os trabalhadores e as políticas e práticas das empresas (BRENKERT, 1992a; MCCALL, 2001; MCMAHON, 1994) e o controle exclusivo das empresas produtivas pelos trabalhadores (DAHL, 1985).

Argumentos para essas estruturas administrativas tomam várias formas. Um tipo de argumento apela para o valor de proteção dos interesses dos trabalhadores (BRENKERT, 1992a; HSIEH, 2005). Um segundo tipo de argumento apela para o valor da autonomia, ou um direito de alguém determinar livremente suas ações, incluindo suas ações no trabalho (MCCALL, 2001). Um terceiro tipo de argumento pela participação dos trabalhadores nas tomadas de decisão administrativas é o argumento do “caso paralelo”. De acordo com ele, se o estado deveria ser governado democraticamente, então as empresas também deveriam, porque empresas são como estados em aspectos relevantes (DAHL, 1985; WALZER, 1983). Um quarto argumento para a participação dos trabalhadores na tomada de decisão nas empresas vê como treinamento valioso, ou mesmo necessário, para a participação em processos políticos na sociedade em geral (COHEN, 1989).

Por considerações de espaço somos impedidos de examinar em detalhes esses argumentos. Mas, críticas geralmente caem em duas categorias. A primeira insiste na prioridade normativa dos acordos, do tipo descrito acima. São poucas as restrições legais sobre os tipos de estruturas administrativas que as empresas podem ter. E algumas empresas são de fato controladas pelos trabalhadores (DOW, 2003; HANSMANN, 1996). Insistir que outras empresas devem ser administradas dessa maneira é dizer, de acordo com esse argumento, que indivíduos não devem ter permissão para organizar suas vidas econômicas como bem entenderem. Outra crítica da participação dos trabalhadores apela para eficiência. Permitir que trabalhadores participem nas tomadas de decisão administrativas pode diminuir o ritmo da tomada de decisão, dado que isso requer dar para muitos trabalhadores a chance de fazer sua voz ser ouvida (HANSMANN, 1996). Isso pode também aumentar os custos de capital da empresa, pois investidores podem demandar termos mais favoráveis se não é dado a eles o controle do empreendimento em troca (MCMAHON, 1994). Ambas as fontes de ineficiência podem colocar a empresa em uma desvantagem significativa em mercados competitivos. E isso pode não ser apenas uma questão de desvantagem competitiva. Se assim fosse, o problema poderia ser resolvido transformando todas as empresas em empresas controladas por trabalhadores. O problema pode ser a diminuição produtiva de forma mais geral.

4. Estruturas usuais para Ética dos Negócios

Eticistas dos negócios procuram entender os contornos éticos e elaborar os princípios da ação correta para a atividade comercial. Uma maneira de avançar esse projeto é pela escolha da estrutura normativa buscando suas implicações para uma série de questões nos negócios.

Uma abordagem influente para ética dos negócios baseia-se na ética das virtudes (por exemplo ALZOLA, 2012; SISON; FONTRONDA, 2012). Moore, em uma série de artigos (BEADLE; MOORE, 2006; G. MOORE, 2005) desenvolve e aplica a ética das virtudes de MacIntyre (1984) aos negócios. Para MacIntyre existem certos bens internos às práticas e certas virtudes são necessárias para alcançá-los. Construindo sobre a tese de MacIntyre, Moore desenvolve a ideia de que os negócios são uma prática e, portanto, há certos bens internos a ela, cuja realização requer o cultivo de virtudes comerciais. Especialistas tem sido inspirado também pela ideia aristotélica de que a vida boa é realizada em comunidade. Eles têm considerado como comunidades dos negócios devem ser estruturadas para ajudar o florescimento de seus membros (HARTMAN, 2015; SOLOMON, 1993).

Outra abordagem de estudo importante da ética dos negócios vem da teoria moral de Kant (ARNOLD; BOWIE, 2003; BOWIE, 1999). A afirmação de Kant de que humanidade deve ser tratada sempre como um fim e nunca somente como um meio tem se provada frutífera, especialmente, por analisar as interações humanas no coração das transações comerciais. Em um mercado competitivo, os indivíduos podem ser tentados a enganar, trapacear ou manipular os outros para ganhar uma vantagem. A teoria moral kantiana destaca essas ações como violações da dignidade humana (SMITH; DUBBINK, 2011).

Teoria ética é, incluindo a teoria das virtudes e a deontologia kantiana, útil para pensar sobre como os indivíduos devem se relacionar uns com os outros no contexto dos negócios (*vide* RORTY, 2006). Mas a ética dos negócios também compreende as leis e as regulamentações que estruturam mercados e organizações. E, aqui, a teoria política parece mais relevante (*vide* MORIARTY, 2005B; PHILIPS; MARGOLIS, 1999). Um número de eticistas dos negócios têm procurado identificar as implicações da justiça como equidade de Rawls (1971) – a teoria da justiça dominante no mundo anglófono – para os negócios. Essa não é uma tarefa fácil, uma vez que, embora Rawls faça algumas observações sugestivas sobre mercados e organizações, ele não articula conclusões específicas nem desenvolve argumentos

detalhados para eles. Mas, especialistas têm argumentado que a justiça como equidade (1) é incompatível com desigualdades significativas de poder e autoridade dentro dos negócios (S. ARNOLD, 2012); (2) requer que as pessoas tenham oportunidade de desempenhar trabalhos significativos (MORIARTY, 2009; HASAN, 2015); e requer formas alternativas de (3) governança corporativa (NORMAN, 2015; SINGER, 2015) e de (4) propriedade corporativa (O'NEILL; WILLIAMSON, 2012).

Uma quarta abordagem para ética dos negócios é a chamada “abordagem de falhas de mercado” (AFM). Uma versão dessa visão pode ser encontrada em McMahon (1981), mas ela tem sido desenvolvida em maiores detalhes e está mais intimamente associada a Heath (2014). De acordo com Heath, a razão pela qual temos uma economia baseada no mercado, em oposição a uma economia de comando, é porque os mercados são mais eficientes. Mas mercados falham, devido a imperfeições das informações, externalidades, custos de transação e assim por diante. O estado corrige muitas falhas de mercado através de regulações. Não estabelecemos limites de poluição e exigimos a verdade nas propagandas, dentre outras coisas. Mas nós não gostaríamos e, não podemos escrever regulamentações para abordar todas as falhas do mercado. É aqui onde entra a ética dos negócios, de acordo com a AFM. Empresários têm uma obrigação moral de não explorar as falhas de mercado que a lei permite que explorem. Em outras palavras, as obrigações morais das pessoas de negócios são identificadas por um regime regulatório ideal – o que teríamos se os regulamentos fossem sem custos e escritos e administrados por uma figura divina.

Selecionar uma estrutura normativa e aplicá-la a uma série de questões é uma forma importante de fazer ética dos negócios. Mas essa não é a única maneira. De fato, a maneira mais comum de abordagem é identificar uma atividade de negócios e, então, analisá-la utilizando intuições e princípios comuns à maioria das teorias políticas e morais. Abaixo, eu abordo questões éticas que surgem no nexos do envolvimento das empresas com três grupos importantes: consumidores, funcionários e sociedade.

5. Empresas e consumidores

A principal maneira de uma empresa interagir com consumidores é através das vendas, ou das tentativas de venda de produtos e serviços para eles. Muitas questões éticas participam dessa interação.

5.1 O que deve ser vendido?

Um número de escritores tem argumentado que algumas coisas não deveriam estar à venda (ANDERSON, 1993; MACDONALD; GAVURA, 2016; SANDEL, 2012; SATZ, 2010). Entre tais coisas que comumente se diz serem inapropriadas para a venda estão serviços sexuais, serviços de barriga de aluguel e órgãos humanos. Alguns autores objetam o mercado desses itens por razões consequencialistas. Eles argumentam que mercados de *commodities* como sexos ou rins levará a exploração de pessoas vulneráveis (SATZ, 2010). Outros objetam as atitudes e valores expressos em tais mercados. Eles afirmam que serviços de barriga de aluguel expressam a atitude que as mulheres são meros receptáculos para a incubação de uma criança (ANDERSON, 1993); mercado de rins sugere que a vida humana pode ser vendida e comprada (SANDEL, 2012), etc.

Outros escritores têm criticado esses argumentos e, em geral, a tentativa de “bloquear” certos bem e serviços do mercado. Brennan e Jaworski (2016) objetam argumentos “expressivistas” contra os mercados de *commodities* polêmicas. Mesmo que vender uma coisa particular por dinheiro expresse desrespeito, eles apontam, isto é culturalmente contingente. Eles e outros também argumentam que os efeitos negativos do mercado acerca das polêmicas *commodities* podem ser eliminados ou, pelo menos, melhorados por meio de regulamentação apropriada e, de qualquer maneira, os bons efeitos desses mercados (por exemplo, uma diminuição no número de pessoas que morrem porque estão esperando por um rim) superam os ruins (ver também TAYLOR, 2005).

5.2 Segurança e confiabilidade do produto

Algumas coisas que empresas podem desejar vender e que as pessoas podem desejar comprar apresentam um risco significativo de dano para o usuário e outros. Quando um produto não é seguro para ser vendido? Essa questão é frequentemente respondida por agências governamentais. Nos Estados Unidos, um número de agências governamentais, incluindo a Comissão de Segurança de Produtos ao Consumidor (CPSC, sigla em inglês), o Sistema Nacional de Segurança de Tráfego de Autoestradas (NHTSA, sigla em inglês), a Administração de Comida

e Medicamentos (FDA, sigla em inglês) são responsáveis por avaliar a segurança dos produtos para o mercado consumidor. Em alguns casos, esses padrões são obrigatórios (por exemplo, medicamentos e dispositivos médicos); em outros casos, são voluntários (por exemplo, trampolins e tendas). O estado identifica padrões mínimos e empresas individuais podem optar por adotar padrões mais altos.

Questões acerca da segurança de produtos são um ponto de debate significativo entre economistas, estudiosos do direito e especialistas em políticas públicas. Os juristas também dedicaram considerável atenção à responsabilidade civil, a área jurídica que lida com casos de danos (não contratuais, nem criminais). Mas os especialistas em ética nos negócios prestam pouca atenção a essas questões. Os tratamentos existentes geralmente combinam discussões de segurança com discussões de responsabilidade – a questão de quem deve pagar pelos danos causados por produtos – e tendem a ser encontradas nos livros de ética dos negócios. Um dos tratamentos mais detalhados é feito por Velasquez (2012). Ele diferencia três visões (compatíveis): (1) a “visão do contrato”, de acordo com a qual o dever do fabricante é divulgar com precisão todos os riscos associados ao produto; (2) a “visão do devido cuidado”, de acordo com a qual o fabricante deve ter o devido cuidado para evitar que os compradores sejam prejudicados pelo produto; (3) a “visão dos custos sociais”, de acordo com a qual o fabricante deve pagar por qualquer prejuízo causado pelo produto, mesmo que o fabricante tenha tomado o devido cuidado para prevenir prejuízo e tenha divulgado com precisão todos os riscos associados com o produto (ver também BROTRIGHT, 2009a).

Há muito espaço para explorar essas questões. Uma área que merece atenção é a de definição de termos chave tais como “segurança” e “risco”. Quedas laterais de berços representam riscos para consumidor, assim como motosserras. Sobre que bases deve ser proibido o primeiro, mas não o segundo (HASNAS, 2010)? Sobre a questão da responsabilidade, um ponto importante é se é justo responsabilizar os fabricantes pelos danos causados por seus produtos, quando os fabricantes não são moralmente culpados por esses danos (PIKER, 1998).

5.3 Publicidade

A maior parte da publicidade contém um componente informativo e um componente persuasivo. A publicidade nos diz algo sobre o produto e tenta nos persuadir a comprá-lo. Ambos os componentes pode ser objetos de avaliação ética.

Enfatizando o componente informacional, alguns autores focam o valor positivo da publicidade. Mercados funcionam eficientemente somente quando certas condições são satisfeitas. Uma dessas condições é informação perfeita: minimamente, consumidores têm que entender as características dos produtos à venda. Conquanto essa condição nunca seja completamente satisfeita na realidade, a publicidade pode ajudar a assegurar que ela será satisfeita em um maior grau (HEATH, 2014). Outro valor que pode ser promovido através da publicidade é autonomia. As pessoas têm certas necessidades e desejos, por exemplo, comer comida saudável, dirigir um carro seguro, os quais suas escolhas como consumidores as ajudam a satisfazer. É mais provável que suas escolhas satisfaçam suas necessidades e desejos se tiverem informações sobre o que está à venda, que a publicidade pode oferecer (GOLDMAN, 1984).

Esses bons efeitos dependem, é claro, das propagandas produzirem crenças verdadeiras ou, pelo menos, não produzir falsas crenças nos consumidores. Escritores tratam isso como uma questão de “engano” em propagandas. A questão aqui não é se a propaganda enganosa é errada – a maioria acredita que é (*vide* CHILD, 1994), mas o que conta como propaganda enganosa e o que a faz errada.

Nos anos 80, a Beech-Nut anunciou como “100% suco de maçã” uma bebida não contendo suco de nenhum tipo. Essas propagandas foram enganosas e, portanto, erradas, porque elas aparentemente fazem afirmações verdadeiras, mas, de fato, fizeram afirmações falsas. Mas muitas propagandas que não parecem enganosas fazem afirmações falsas ou não verificáveis. Considere o slogan da Calvin Klein “entre o amor e a loucura encontra-se a obsessão” ou o slogan da Gillette, “o melhor que um homem pode ter”. É comum dizer que esses tipos de afirmações não estão comprometidos com a verdade e, então, não podem ser enganosas (CARSON, 2010). Ainda que essas afirmações possam, de fato, enganar algumas pessoas. Um indivíduo pode acreditar na afirmação da Gillette de que seu produto é o melhor que um homem pode ter e comprá-lo sob essa base. Os reguladores resolvem essa complicação empregando um padrão de “pessoa razoável” (ATTAS, 1999). As propagandas são consideradas enganosas quando uma pessoa razoável, e não uma pessoa qualquer, afinal de contas, é enganada. Isso torna a publicidade enganosa uma questão de resultados para os consumidores, não intenções para os anunciantes.

Muitas razões têm sido oferecidas para o “porquê” de a propaganda enganosa ser errada. Uma é a afirmação kantiana de que enganar os outros é desrespeitá-los, um uso deles como meros meios. A propaganda enganosa pode

também levar dano aos consumidores (quem comprou produtos sub-ótimos, dado seus desejos) e competidores (que perderam vendas). Uma crítica final à propaganda enganosa é que ela erode a confiança em sociedade (ATTAS, 1999). Quando as pessoas não confiam umas nas outras, elas não irão se envolver em transações econômicas, ou se envolverão apenas com caras proteções legais.

O componente persuasivo da publicidade também é um assunto frutífero da investigação ética. Galbraith (1958), um dos primeiros críticos, acha que a publicidade, em geral, não informa às pessoas como adquirir o que elas desejam, mas, ao invés disso, dá a elas novos desejos. Ele chama isso de “efeito dependência”: nossos desejos dependem do que é produzido, não **vice-versa**. Além disso, dado que nós somos inundados por publicidade de bens de consumo, nós desejamos muito deles e não desejamos suficientemente bens público. Hayek (1961) rejeita essa defesa argumentando que pouco, se algum, dos nossos desejos são independentes do nosso ambiente e, que de qualquer maneira, desejos produzidos em nós por publicidade não são menos significativos que os desejos produzidos em nós de outras maneiras.

Galbraith está preocupado sobre os **efeitos** persuasivos das propagandas. Em contratos, escritores recentes focam-se nas **técnicas** de persuasão que os anunciantes usam. Alega-se que algumas delas cruzam a linha de manipulação. É difícil definir manipulação precisamente, embora muitas tentativas hajam sido feitas (por exemplo COONS; WEBER, 2012). Para os nossos propósitos, propaganda manipulativa pode ser entendida como uma propaganda que tenta persuadir consumidores, frequentemente (mas não necessariamente), utilizando-se de meios não-rationais, a fazerem escolhas irracionais ou sub-ótimas, dadas suas próprias necessidades e desejos (*vide* BEAUCHAMP, HARE; BIEDERMAN, 1984; BRENKERT, 2008).

A publicidade associativa é frequentemente considerada um exemplo de publicidade manipulativa. Em publicidade associativa, o anunciante tenta associar um produto com uma crença, sentimento, atitude ou atividade positiva, as quais têm pouco a ver com o próprio produto. Assim, muitos comerciais televisivos de caminhões nos Estados Unidos associam caminhões à masculinidade. Comerciais para perfumes associam estes produtos com sexo entre pessoas bonitas. A sugestão é que se você é um certo tipo de pessoa (por exemplo, viril), então terá um certo tipo de produto (por exemplo, um caminhão). Crisp (1987) argumenta que esse tipo de publicidade objetiva criar desejos nas pessoas contornando sua faculdade de escolha consciente, e ao fazer isso, subverte sua autonomia (cf. Arrington, 1982; Phillips, 1994). Lippke (1989) argumenta que isso faz as pessoas desejarem as

coisas erradas, encorajando-nos a tentar satisfazer nossos desejos não mercantis (por exemplo, ser mais viril) através de meios mercantis (por exemplo, comprando um caminhão). Quão sério nós tomamos essas críticas pode depender de quão efetivo nós pensamos serem as propagandas associativas e outras formas. Na medida em que achamos que os anunciantes não conseguem “contornar” nossa faculdade de escolha consciente, podemos ficar menos preocupados e mais espirituosos com suas tentativas de fazê-lo.

Nossos julgamentos sobre essas questões podem ser sensíveis ao contexto. Enquanto a maioria das pessoas podem ser capazes de ver através das tentativas das propagandas para persuadi-las, algumas podem não ser (pelo menos em algum momento). Paine (PAINE *et al.*, 1984) argumenta que propaganda é justificada porque ela ajuda os consumidores a tomarem decisões sábias no mercado. Mas crianças, ela argumenta, não possuem a capacidade de fazer escolhas de consumo sábias (ver também MOORE, 2004). Então, propagandas direcionadas para crianças – ao contrário de propagandas de produtos para crianças direcionadas a adultos – constituem uma forma de exploração objetável. Outras populações que podem ser similarmente vulneráveis são os senis, os ignorantes e os enlutados. A ética pode requerer não um total banimento do *marketing* para eles, mas um cuidado especial em como eles são comercializados para tais grupos (BRENKERT, 2008).

5.4 Vendas e Negociações

Vendas são centrais para os negócios. Talvez, surpreendentemente, eticistas dos negócios têm dito pouco diretamente sobre vendas. Mas muito do que é dito sobre propaganda é também aplicado as vendas. Vendedores são, em um sentido, o anunciante final dos produtos para os consumidores. Eles fornecem benefícios aos consumidores da mesma maneira que os anunciantes e têm a mesma capacidade de enganar ou manipular os consumidores.

Carson (2010) desenvolveu uma teoria ética detalhada para vendedores. De acordo com ele, vendedores têm, pelo menos, os seguintes quatro deveres *pro tanto*: (1) fornecer aos clientes avisos e precauções de segurança; (2) abster-se de mentir e enganar; (3) responder completamente as questões dos consumidores sobre os itens; (4) não direcionar os clientes para compras inadequadas para eles, dadas as necessidades e desejos declarados. Carson justifica (1)-(4) apelando para a regra de ouro: trate os outros como você deseja ser tratado. Ele identifica dois

outros deveres que os vendedores **podem** ter (ele é agnóstico): (5) não venda produtos ao cliente que você (o vendedor) julga inadequados para eles, dadas as necessidades e desejos deles, sem dizer aos clientes por que você acha isso; e (6) não venda aos clientes produtos de baixa qualidade ou defeituosos, sem dizer a eles por que você pensa assim. Para a maior parte, (1) – (4) pedem ao vendedor para não prejudicar o consumidor; (5) e (6) pedem ao vendedor para ajudar o consumidor, em particular, ajudá-lo a não cometer erros tolos. Se vendedores devem ajudar os consumidores dessa maneira pode depender de quão adverso o relacionamento deles deve ser.

A questão mais ampla aqui é de divulgação. Holley (1998) argumenta que os vendedores devem divulgar para clientes o que uma “pessoa razoável gostaria de saber” sobre um produto antes de comprá-lo. Ebejer e Morden (1988) afirmam que vendedores devem divulgar todas as informações “relevantes” para a aquisição por um comprador. Mas não há consenso sobre que informação é relevante para uma decisão de compra, ou o que uma pessoa razoável deseja saber.

Para muitos produtos comprados e vendidos nos mercados, vendedores oferecem um item a um certo preço, e compradores aceitam ou não esses preços. Mas em alguns casos há negociação sobre o preço (e sobre outros aspectos da transação). Vemos isso na venda de itens “caros”, como carros e casas, e em salários para empregos. Embora existam muitas questões éticas que surgem na negociação, uma questão que recebeu atenção especial é a de “blefar” ou deliberadamente distorcer a posição de barganha de alguém. O *locus classicus* desse debate é Carr (1968). De acordo com ele, blefar em uma negociação é permissível porque os negócios possuem seu próprio conjunto especial de regras e blefar é permitido de acordo com essas regras. Na visão de Carr, todos que entram na arena dos negócios aceitam blefar como permissível, assim como todos que entram em um ringue de box aceitam socar as pessoas como permissível. Carson (2010) concorda que blefar nos negócios é permissível, embora em um escopo mais limitado de casos do que Carr aceita. A defesa de Carson para o ato de blefar apela para autodefesa. Se você tem boas razões para acreditar que seu adversário em uma negociação está distorcendo a posição de barganha dele, então você tem permissão para distorcer a sua. Uma exigência para falar a verdade nessas circunstâncias nos colocaria em uma desvantagem significativa em relação ao nosso adversário, a qual você não é obrigado a sofrer. Uma implicação da visão de Carson é que você não tem permissão para distorcer sua posição de barganha se você não tem boas razões para acreditar que seu adversário está distorcendo a dele.

5.5 Atribuição de preços

Em modelos simplificados de mercado, compradores e vendedores individuais são “tomadores de preços”, não “fazedores de preços”. Isto é, os preços dos bens e serviços são estabelecidos pelas forças agregadas da oferta e demanda; nenhum indivíduo é capaz de comprar e vender um bem por qualquer outra coisa que não o preço de mercado. Na realidade, as coisas são diferentes. Vendedores de bens têm alguma flexibilidade sobre como atribuir preço às mercadorias.

A maioria dos especialistas em ética dos negócios aceita que, na maioria dos casos, o preço pelo qual os produtos devem ser vendidos é uma questão para indivíduos privados decidirem. Essa visão tem sido defendida sob o fundamento dos direitos de propriedade. Da afirmação de que se eu tenho um direito a X, então eu sou livre para transferi-lo para você sob quaisquer que sejam os termos que eu proponho e você aceita (BOATRIGT, 2010). Isso tem sido defendido sobre os fundamentos do bem-estar. Preços estabelecidos por trocas voluntárias de indivíduos revelam informações valiosas sobre a demanda relativa a respeito do suprimento de um bem, permitindo que os recursos fluam em usos mais produtivos (HAYEK, 1945). Apesar disso, a maior parte dos especialistas em ética reconhece alguns limites na atribuição de preço.

Uma questão que recebeu atenção recentemente é a discriminação de preços. Com isso, não quero dizer discriminação nos preços com base na associação de uma pessoa a classes tradicionalmente protegidas, como raça e sexo (isso é amplamente considerado errado). “Discriminação de preço” refere-se à discriminação de preços com base na disposição das pessoas a pagar: cobrar mais das pessoas que estão dispostas a pagar mais e cobrar menos das pessoas que estão dispostas a pagar menos. Economistas tendem a pensar que a discriminação de preços é valiosa na medida em que permite que as empresas aumentem a produção. Mas o estatuto moral dessa prática é menos claro. Quando foi revelado que a Staples e outros varejistas on-line estavam cobrando dos consumidores em diferentes códigos postais preços diferentes pelos mesmos produtos ao mesmo tempo, os consumidores ficaram indignados. Mas alguns autores argumentam que essa prática não pior do que os cinemas dando descontos para crianças (ELEGIDO, 2011; MARCOUX, 2006). O problema pode ser que a Staples e outras empresas se engajam nessa prática sem o divulgar. Ao fazer isso, elas estariam tirando vantagem da ignorância dos consumidores.

Outra questão ética do estabelecimento de preços é a manipulação deles. Manipulação de preços pode ser entendida como um aumento acentuado no preço de um bem necessário após uma emergência que torna esse bem escasso. Logo após o impacto do Furacão Katrina em Nova Orleans em 2005, por exemplo, muitos varejistas cobraram preços muito altos pela água e pela gasolina. Muitas jurisdições possuem leis sobre a manipulação de preços e essa prática é amplamente tomada como não ética (SNYDER, 2009). A razão para isso é que é um caso clássico de exploração: *A* extrai uma taxa excessiva de *B* em circunstâncias nas quais *B* não pode razoavelmente recusar a oferta de *A* (VALDMAN, 2009). Mas alguns teóricos defendem a manipulação de preços. Enquanto tomam como garantido que as vendas de itens em circunstâncias como aquelas são exploração, eles notam que elas são mutuamente benéficas. Tanto o vendedor quanto o comprador preferem se engajar nessas transações ao invés de não o fazer. Além disso, quando itens são vendidos por preços inflacionados, eles atraem mais vendedores para o mercado. Permitir manipulação de preços pode então ser a maneira mais rápida para eliminá-la (ZWOLINSKI, 2008) (para mais discussões, ver a entrada da SEP sobre **exploração**).

Muitos especialistas contemporâneos acreditam que vendedores têm um amplo, embora não ilimitado, controle sobre quanto podem cobrar por bens e serviços. Mas há uma tradição antiga em ética dos negócios, encontrada em Aquino e outros pensadores medievais, de acordo com a qual há um preço que o vendedor deve cobrar: o “preço justo”. Há o debate sobre o que exatamente os pensadores medievais queriam dizer por “preço justo” (REIFF, 2013). De acordo com a interpretação historicamente comum, o preço justo é determinado pelos custos de produção do vendedor, i.e., o preço que compensa o vendedor pelo valor de seus gastos e trabalho. Interpretações mais recentes entendem o preço justo medieval como algo próximo do preço de mercado, o qual pode ser mais ou menos que o preço de produção (KOEHN; WILBRATTE, 2012). No tratamento contemporâneo, Reiff ressuscita e reabilita o conceito de preço justo, distinguindo-o do preço de mercado e especifica-o como “o custo social médio total do bem em questão” (2013, p.109). Uma visão com a de Reiff deve lidar com os argumentos esboçados acima – baseados em direitos e bem-estar – para permitir que os preços sejam determinados pelas escolhas voluntárias de compradores e vendedores.

6. Empresas e trabalhadores

Especialistas em ética dos negócios tem escrito muito sobre o relacionamento entre empregadores e empregados. Muitos desses escritos têm investigado as obrigações que os empregadores devem aos funcionários. Isso pode ser porque empregadores usualmente tem mais poder que os funcionários e, portanto, têm maior discricionariedade na maneira como tratam os funcionários do que os empregados na maneira como tratam os empregadores. Abaixo, identifiquei quatro questões acerca da interface empregado/empregado: (1) contratação e demissão; (2) pagamento; (3) trabalho significativo e (4) informantes (*whistleblowing*). Outro tópico importante nessa área é privacidade. Por razões de espaço, esse tópico não será discutido, mas veja as entradas sobre **privacidade** e **privacidade e tecnologia da informação** da SEP.

6.1 Contratação e demissão

Problemas éticos acerca da contratação e demissão focam a questão: Qual critério o empregador deve utilizar, ou não utilizar, em decisões de emprego? A questão de qual critério empregadores não deveriam utilizar é abarcada nas discussões acerca da discriminação.

Enquanto há algum debate sobre se a discriminação nos empregos deve ser legalmente proibida (*vide* EPSTEIN, 1992), quase todos concordam que isso é moralmente errado (HELLMAN, 2008; LIPPERT-RASMUSSEN, 2014). As discussões têm se focado em duas questões: Primeiro, quando o uso de certos critérios em uma decisão de empregos conta como discriminatório? Parece errado que o Wal-Mart exclua candidatos brancos para um trabalho em seu departamento de marketing, mas não parece errado para a Hovey Playes (uma trupe de teatro) excluir candidatos brancos para uma produção de *A Raisin in the Sun*. Nós podemos dizer que se uma prática de contratação é discriminatória dependendo de se o critério usado é relevante para o trabalho. Mas isso pode não ir longe o suficiente, como revela o caso das “qualificações de reação”. Suponha que comensais brancos preferem ser servidos por garçons brancos ao invés de garçons negros. Nesse caso a raça parece ser relevante para o trabalho, mas também parece errado para os empregadores considerarem raça. Outra questão que tem recebido considerável atenção é: o que faz a discriminação errada? Alguns argumentam que discriminação é errada por

seus efeitos sobre aqueles que são discriminados (LIPPERT-RASMUSSEN, 2014), outros pensam que isso é errado por causa do que ela expressa (HELLMAN, 2008). Para mais discussões, consulte o verbete sobre **discriminação**, na SEP.

Alguns autores acreditam que as obrigações dos empregadores não são satisfeitas simplesmente quando eles **evitam** usar certos critérios nas decisões de contratação. De acordo com eles, empregadores tem o dever de contratar o candidato mais qualificado. Alguns justificam esse dever apelando para considerações de mérito (MILLER, 1999); outros justificam apelando para a igualdade de oportunidade (MASON, 2006). O desafio padrão para essa visão apela para o direito de propriedade (KERSHNER, 2004). Uma oferta de emprego tipicamente implica uma promessa de pagamento de uma soma de seu dinheiro ao candidato do emprego pela execução de certas tarefas. Embora possamos pensar que excluir algumas maneiras pelas quais você pode dispor de sua propriedade (por exemplo, regras contra discriminação na contratação) pode ser justificado, podemos pensar que excluir todas as maneiras, exceto uma (a saber, um requisito para contratar o candidato mais qualificado) é injustificado. Em apoio a isso, nós podemos pensar que a proprietária de um pequeno negócio não faz nada de errado quando ela contrata sua filha para um trabalho de meio período ao contrário de contratar um estranho mais qualificado.

Muitos dos mesmos problemas éticos levantados pela contratação também são levantados pela demissão. Tem havido uma discussão robusta sobre ética da demissão na literatura de ética dos negócios. Há dois campos principais: aqueles que pensam que o emprego deve ser “discricionário”, para que um empregador possa demitir um funcionário por qualquer motivo (EPSTEIN, 1984; MAITLAND, 1989), e aqueles que pensam que os empregadores devem ser capazes de demitir um funcionário apenas por “justa causa” (e.g., mau desempenho ou crise nos negócios) (MCCALL; WERHANE, 2010). De fato, poucos escritores defendem uma versão “pura” da visão “discricionária”. A maioria diria que é errado que um empregador demita um funcionário por algumas razões, e.g., a descoberta que ele é muçulmano ou sua recusa em cometer um crime pelo empregador. Então o debate é entre aqueles que pensam que o empregador deve ser capaz de demitir um funcionário por qualquer razão **com algumas exceções**, e aqueles que pensam que os empregadores devem ser capazes de demitir funcionários apenas por certas razões. Nos EUA, a maioria dos funcionários está sob o regime discricionário, enquanto na Europa, a maioria dos funcionários é protegida, após um período de estágio, por algo análogo à justa causa. Argumentos pela justa causa apelam para os efeitos que a demissão tem sobre os funcionários individuais, especialmente aqueles que

tem trabalhado para um empregador por muitos anos (MCCALL; WERHANE, 2010). Argumentos em favor da visão discricionária apelam para a liberdade ou efeitos macroeconômicos. Alega-se, no primeiro caso, que justa causa é uma restrição injustificada à liberdade de empregadores e empregados (EPSTEIN, 1984) e, no último caso, que a discricionariedade aumenta a taxa de desemprego (MAITLAND, 1989).

6.2 Pagamento

As organizações empresariais geram receita e parte dessa receita é distribuída aos seus funcionários na forma de pagamento. Como a demanda por pagamento normalmente excede a oferta, a questão de como o pagamento deve ser distribuído é naturalmente analisada como um problema de justiça.

Dois teorias da justiça gerais acerca do pagamento têm atraído atenção. Uma pode ser chamada de “visão do acordo”. Segundo essa visão, o salário justo é o salário que empregador e empregado concordaram sem uso da força ou fraude (BROATRRIGHT, 2010; MACHAN; CHESHER, 2002). Essa visão é justificada algumas vezes em termos de direitos de propriedade. Funcionários possuem sua força de trabalho e empregadores possuem seu capital, e eles são livres, dentro de limites amplos, para dispor de suas propriedades como eles quiserem (BROATRRIGHT, 2010). Uma segunda visão acerca dos salários pode ser chamada de “visão da contribuição”. De acordo com ela, o salário justo para um trabalhador é o salário que reflete sua contribuição para a empresa. Essa visão vem em duas versões. Sobre a versão absoluta, trabalhadores devem receber uma quantia de pagamento que equivale ao valor das suas contribuições para a empresa (MILLER, 1999). Sobre a versão comparativa, trabalhadores devem receber uma quantia de pagamento que reflita o valor relativo as suas contribuições para a empresa, dado o que os outros na empresa contribuem e são pagos (STERNBERG, 2000). A visão da contribuição parece ser normativamente básica, uma visão pela qual nenhum outro argumento pode ser dado (MILLER, 1999). Uma analogia com a punição pode ser estabelecida: assim como parece intuitivamente correto a severidade da punição de um crime refletir a gravidade de seu crime, então pareceria intuitivamente correto o valor do pagamento do trabalhador refletir o valor de seu trabalho.

O pagamento de qualquer funcionário em uma empresa pode ser avaliado do ponto de vista moral usando as duas teorias esboçadas acima. Mas, especialistas em ética dos negócios dão atenção particular para o pagamento de certos grupos

de funcionários, como CEOs e trabalhadores em condições precárias (*sweatshops*).

Tem havido um robusto debate sobre se CEOs estão recebendo demais (MORIARTY, 2005a), com os estudiosos se dividindo, *grosso modo*, em dois campos. Aqueles no campo do “poder de gestão” acreditam que CEOs exercem poder sobre conselhos de administração e usam esse poder para extrair rendas de suas empresas acima do mercado. Aqueles no campo da “contratação eficiente” acreditam que as negociações acerca de pagamentos entre os CEOs e os conselhos são geralmente realizadas em condições de igualdade e os grandes pacotes de remuneração dos CEOs refletem suas habilidades raras e valiosas (*vide* REIFF, 2013).

Também tem havido um grande debate se trabalhadores em condições precárias estão recebendo tão pouco. Alguns dizem ‘não’ (POWELL; ZWOLINSKI, 2012; ZWOLINSKI, 2007). Eles dizem que salários de *sweatshops*, embora baixos para os padrões de países industrializados, não são baixos para os padrões dos países nos quais tais empresas estão localizadas. Isso explica por que as pessoas escolhem trabalhar em uma *sweatshop*: é a melhor oferta que eles têm. Esforços para aumentar artificialmente os salários de trabalhadores de *sweatshop*, de acordo com esses escritores, é um equívoco sob duas acusações. Primeiro, é uma interferência nas escolhas autônomas dos empregadores e trabalhadores. Segundo, isso provavelmente faz que os trabalhadores fiquem em pior posição, dado que os empregadores responderão ou movendo suas operações para um novo local ou empregando poucos trabalhadores em sua localização. Outros especialistas desafiam essa defesa. Embora concedam que os trabalhadores optem por trabalhar em *sweatshops*, eles negam que suas escolhas sejam voluntárias (ARNOLD; BOWIE, 2003). Dados os salários muito baixos deles, isso sugere que trabalhadores de *sweatshops* são explorados. Além disso, alguns argumentam, apelando para um dever kantiano de beneficência, de que empresas podem e devem fazer mais pelos trabalhadores de *sweatshop* (SNYDER, 2010). Em resposta a afirmação de que as empresas se colocam em uma desvantagem competitiva se elas fizerem mais pelos trabalhadores, escritores tem apontado para casos atuais onde as empresas têm sido capazes de assegurar um melhor tratamento para trabalhadores de *sweatshop* sem sofrerem grande penalidade financeiras (HARTMAN, ARNOLD; WOKUTCH, 2003).

6.3 Trabalho Significativo

Smith, em 1776 (1976), observou, e é amplamente aceito, que uma divisão detalhada do trabalho aumenta muito a produtividade dos processos de manufatura. Para usar o exemplo dele: se um trabalhador executa todas as tarefas necessárias para fazer um alfinete, ele pode fazer apenas alguns deles por dia. No entanto, se o trabalhador se especializar em uma ou duas dessas tarefas e combinar seus esforços com outros trabalhadores que se especializarem em uma ou duas outras tarefas, então, juntos eles podem fazer centenas de alfinetes por dia. Mas há um custo humano da divisão detalhada do trabalho, de acordo com Smith. Executar uma ou duas tarefas simples todos os dias provavelmente faz um trabalhador “tão estúpido e ignorante quanto é possível para uma criatura humana se tornar” (SMITH, 1976, v.1, p. 178).

Apelar para “trabalhos significativos” são uma resposta a esse problema. A implicação de um apelo a um trabalho significativo não é um apelo para que o trabalho seja mais “importante”, isto é, para contribuir para a produção de um bem ou serviço que seja objetivamente valioso ou que os trabalhadores acreditem ser valioso. Ao invés disso, é apelar para que os processos de trabalho sejam organizados de tal maneira que o trabalho seja interessante, requeira habilidades e dê aos trabalhadores poderes substanciais de tomada de decisão (ARNESON, 1987; MICHAELSON *et. al.* 2014).

A ideia de Smith que os processos de trabalhos são mais eficientes quando são divididos em segmentos sem sentido, leva alguns autores a acreditar que em uma economia competitiva, empresas não irão oferecer tantos trabalhos significativos quanto os trabalhadores desejam (WERHANE, 1985). Em resposta, tem-se argumentado que há um mercado de trabalho e se os trabalhadores desejam trabalhos significativos, então os empregadores teriam um incentivo para providenciá-los (MAITLAND, 1989; NOZICK, 1974). De acordo com esse argumento, na medida em que vemos “tão poucos” trabalhos significativos em oferta, isso é porque os trabalhadores não os preferem – ou, mais precisamente, porque trabalhadores estão dispostos a trocar mais **significado** por outros benefícios, tais como altos salários. Esse argumento assume, é claro, que trabalhadores têm a habilidade financeira para trocar salários por significado pessoal.

O argumento acima trata trabalho significativo como uma questão de preferência: como uma comodidade que os empregadores podem se recusar a oferecer ou que os trabalhadores podem trocar. Outros resistem a esse entendimento.

De acordo com Schwartz (1982), os empregadores devem oferecer aos empregados um trabalho significativo e os funcionários devem executá-los, por respeito à autonomia. O raciocínio é: as pessoas autônomas fazem suas próprias escolhas; elas não seguem sem consideração as decisões dos outros. Uma dificuldade para esse argumento é que o respeito à autonomia não parece exigir que façamos **todas** as escolhas por nós mesmos. Uma pessoa pode, ao que parece, escolher autonomamente ao permitir que decisões importantes sejam tomadas no lugar dela em certas esferas de sua vida, por exemplo, por um treinador, por um membro da família ou um comandante militar.

Um problema potencial para essa resposta nos leva de volta a Smith e aos argumentos “formativos” para o trabalho significativo. O problema, de acordo com alguns autores, é que, se a maior parte do dia de uma pessoa é cedido para atividades sem significado, então sua capacidade de escolha autônoma, e talvez suas outras faculdades intelectuais, podem se deteriorar. Um apelo por trabalho significativo pode então ser entendido como um apelo por organização de locais de trabalhos para que essa deterioração não ocorra (ARNESON, 2009; S. ARNOLD, 2012). Além de Smith, Marx, em 1844 (2000) estava claramente preocupado com os efeitos do trabalho no florescimento humano. Argumentos formativos enfrentam duas dificuldades. Uma é estabelecer a conexão entre trabalho sem significado e escolhas autônomas (ou outra faculdade intelectual). Segundo, e talvez mais importante, argumentos formativos fazem certas suposições acerca da natureza do bem e do papel público de promovê-lo. Eles assumem que é melhor para as pessoas terem suas faculdades de escolha autônomas completamente desenvolvidas (etc.) e que o público deve ajudar a desenvolvê-las. Ambas as suposições podem ser desafiadas. Neutralistas em filosofia política pensam que o estado não deveria promover o bem, pelo menos quando há um desacordo razoável sobre o que é o bem (por exemplo RAWLS, 1993).

6.4 Informantes

Suponha que você descubra que sua empresa está pagando suborno para funcionários públicos para ganhar contratos. Para parar isso, uma coisa que você

pode fazer é “botar a boca no trombone” (*whistleblowing*)⁴¹, denunciando essa informação para um jornalista. Embora diferentes teóricos deem diferentes definições para a atividade do informante (por exemplo BRENKERT, 2010; DAVIS, 2003; DEGEORGE, 2009), os seguintes elementos estão frequentemente presentes: (1) *status* privilegiado; (2) informação não pública; (3) atividade ilegal ou imoral, (4) evitar a cadeia de comando usual na empresa, (5) intenção de resolver um problema. No exemplo acima, você seria um denunciante porque você é (1) um funcionário (2) que divulga informação não pública (3) acerca de uma atividade ilegal na empresa (4) para pessoas de fora dela (5) em um esforço para parar essa atividade.

Debates sobre informantes tendem a focar a questão de quando a denúncia é justificada – no sentido de quando ela é permitida ou quando ela é requerida. Esse debate assume que a denúncia requer justificação, ou é errada, mantendo-se iguais outros fatores. Muitos especialistas em ética dos negócios fazem essa suposição sob o fundamento de que funcionários tem um dever *pro tanto* de lealdade a suas empresas (por exemplo BOATRIGHT, 2009a). Contra isso, alguns argumentam que a relação entre a empresa e o funcionário é puramente transacional, uma troca de dinheiro por trabalho (DUSKA, 2000) e, sendo assim, não é normativamente robusta o suficiente para fundamentar um dever de lealdade. Para uma discussão desse tópico, confira o verbete **lealdade** da SEP.

Uma justificação proeminente para denúncias é devida a DeGeorge (2009). De acordo com ele, é permissível para um funcionário denunciar quando sua ação prevenir danos a sociedade. Em uma descrição similar, Brenkert (2010) diz que o dever de denunciar deriva do dever de prevenir atos errados. O dever de prevenir dano tem um peso maior que o dever de lealdade. Para determinar se denunciar é não apenas permissível, mas obrigatório, DeGeorge diz, nós devemos levar em consideração o provável sucesso da denúncia e seus efeitos sobre o próprio denunciante. Humanos são criaturas tribais, e denunciadores são, frequentemente, tratados mal por seus colegas. Então, se é improvável que a denúncia seja bem-

⁴¹ N.T.: A expressão idiomática *whistleblowing* (literalmente, “soprar apito”) não tem tradução exata, razão por que optamos por “botar a boca no trombone”, que cumpre a mesma função na linguagem popular e decorre de uma metáfora similar (ou seja, o ruído decorrente de um instrumento de sopro). No entanto, no decorrer do texto, traduzimos a ação correspondente por “denúncia” e o respectivo agente por “informante” – termos usualmente empregados neste contexto.

sucedida, então não é necessário tentar. A falta de uma exigência moral de denunciar nestes casos pode ser vista como uma instância específica da regra de que indivíduos não precisam fazer grandes sacrifícios pessoais para promover os interesses dos outros, mesmo quando esses interesses são importantes.

Outra descrição de denúncia é dada por Davis (2003). Assim como Brenkerf (e diferente de DeGeorge), Davis foca as ações erradas em que a empresa se engaja (não nos danos que isso causa). De acordo com Davis, no entanto, o objetivo de denunciar não é tanto impedir o erro, mas evitar a própria cumplicidade. Ele diz que uma empregada é obrigada a denunciar sua empresa quando ela acredita que está envolvida em um comportamento seriamente injusto, e seu trabalho para a empresa “contribuirá ... para o erro, se ... [ela] [não] revelar publicamente o que [ela sabe]” (2003, p.550). Uma característica não usual da descrição de Davis é que ela limita os denunciantes a pessoas que estão atualmente dentro das empresas. Você não pode “botar a boca no trombone” a respeito de uma empresa depois que você a tiver deixado.

O informante revela um fenômeno real e importante. Mas isso não parece moralmente distinto, no sentido de que os valores e os deveres envolvidos aí são familiares. Lealdade a um indivíduo (ou grupo) pode requerer que nós demos preferência a seu (ou seus) interesses, até certo ponto. E ainda, em geral, nós devemos evitar cumplicidade com comportamentos imorais e devemos também fazer um esforço para prevenir danos e ações erradas, especialmente quando nossos esforços são, provavelmente, bem-sucedidos e não são muito custosos para nós. Dada as descrições apresentadas acima, denunciar é simplesmente tentar agir de acordo com esses valores e executar esses deveres no contexto do local de trabalho.

7. A empresa na sociedade

Empresas como um todo comandam recursos enormes e como resultado podem ter um enorme impacto na sociedade. Uma maneira como elas impactam na sociedade, é claro, é produzindo bens e serviços e provendo empregos. Mas elas podem também impactar a sociedade por tentar resolver problemas sociais e por utilizar seus recursos para influenciar regulamentações e leis do estado.

7.1 Responsabilidade social corporativa

“Responsabilidade social corporativa”, ou RSC, é tipicamente entendida como ações de empresas que são (i) não legalmente exigidas, e (ii) objetivam beneficiar outras partes que não a empresa (sendo os benefícios das empresas entendidos em termos de retorno sobre patrimônio, retorno sobre ativos ou alguma outra medida de desempenho financeiro). As partes beneficiadas podem ser mais ou menos associadas de modo próximo à própria empresa; elas podem ser os funcionários da própria empresa ou pessoas em lugares distantes.

Um exemplo famoso de RSC envolve a companhia farmacêutica Merck. Nos anos 70, Merck estava desenvolvendo uma droga para o tratamento de parasitas no gado e foi descoberto que uma versão da droga poderia ser usada para tratar Cegueira do Rio, uma doença que causa coceira debilitante, dor e, eventualmente, cegueira. O problema foi que a droga custaria milhões de dólares para ser desenvolvida, e geraria pouco ou nenhum dividendo para a Merck, dado que pessoas afligidas pela Cegueira do Rio – milhões de africanos subsaarianos – eram muito pobres para pagar por ela. No fim, a Merck decidiu por desenvolver a droga. Como esperado, ela foi efetiva para o tratamento da Cegueira do Rio, mas a Merck não obteve lucro com ela. Até o momento em que este artigo foi escrito em 2016, a Merck, agora em conjunto com várias organizações não-governamentais, continua fabricando e distribuindo o medicamento gratuitamente em todo o mundo em desenvolvimento.

Responsabilidade social corporativa, ou RSC, não é somente o termo usado pelos especialistas em ética dos negócios para descrever ações como a da Merck. Elas também podem ser descritas como um exemplo de “cidadania corporativa” ou “sustentabilidade corporativa” (CRANE; MATTEN; MOON, 2008; NÉRON; NORMAN, 2008). É duvidoso que algo importante esteja na escolha de rótulos.

A literatura especializada sobre RSC é dominada por cientistas sociais. A questão deles é tipicamente se, quando e como ações socialmente responsáveis beneficiam a empresa financeiramente. A sabedoria convencional parece ser a de que existe uma ligeira correlação positiva entre desempenho social corporativo e desempenho financeiro corporativo, mas não está claro para que lado a causalidade ocorre (MARGOLIS; WALSH, 2003; ORLITZKY *et al.*, 2003; VOGEL, 2005). Isso é, não é claro se um comportamento pró-social por parte das empresas causa a elas sua recompensa financeira (por exemplo, por consumidores que valorizam seu comportamento), ou se o sucesso financeiro causa na empresa o engajamento em mais comportamentos pró-sociais (por exemplo, liberando recursos que de outra

forma seriam gastos no núcleo das funções empresariais). Já que nossa preocupação é com questões normativas, nós focaremos nas razões morais a favor e contra RSC.

Alguns autores conectam o debate sobre RSC com o debate sobre os fins da governança corporativa. Assim, Friedman (1970) objeta a RSC dizendo que os administradores deveriam maximizar a riqueza dos acionistas. Acredita-se que a teoria das partes interessadas seja mais adaptável à atividade pró-social das empresas, pois permite que elas façam outras coisas além de aumentar a riqueza dos acionistas. Mas nós não precisamos ver o debate acerca da RSC como argumentos sobre os fins apropriados da governança corporativa. Nós podemos vê-lo como um debate acerca dos meios para aqueles fins, com alguns afirmando, e outros negando, que certos atos de comportamento pró-social são obrigatórios não importando quais fins uma empresa persegue.

Muitos escritores dão uma razão consequentialista ampla para RSC. O argumento tende a ir como se segue: (i) há sérios problemas no mundo tais como, pobreza, conflitos, degradação ambiental e etc.; (2) qualquer agente com recursos e conhecimento necessários para melhorar esses problemas tem uma responsabilidade moral de fazê-lo, assumindo que os custos para ele mesmo não são grandes; (3) empresas têm recursos e conhecimento necessários para melhorar esses problemas sem incorrer em grandes custos; portanto, (4) empresas devem melhorar esses problemas. A visão de que **alguém** deve fazer **algo** acerca dos problemas do mundo parece claramente verdade para muitas pessoas. Não é somente uma oportunidade de aumentar o bem-estar social aliviando o sofrimento, pessoas em sofrimento também podem ter direito à assistência. A questão controversa é quem deve fazer alguma coisa para ajudar, e quanto se deve fazer. Então defensores do argumento acima focam mais suas atenções em estabelecer que **empresas** tem esses deveres, contra aqueles que dizem que esses deveres são apropriadamente atribuídos ao estado ou indivíduos. O'Neill (2011) e Wettstein (2009) argumentam que empresas são "agentes da justiça", assim como estados e indivíduos e têm deveres de ajudar os necessitados. Strudler (2017) legitima o comportamento altruísta das empresas, minando a alegação de que os acionistas são os proprietários e, portanto, lhes é devida a riqueza excedente. Hsieh (2004) diz que, mesmo se admitirmos que as empresas não têm obrigações sociais, os indivíduos as têm, e a melhor maneira de muitas pessoas cumpri-las é através das atividades de suas empresas (ver também McMahon, 2013).

Debates sobre as RSC não são apenas debates sobre se males sociais específicos devem ser tratados por empresas específicas. Eles são também debates

acerca de em que tipo de sociedade nós desejamos viver. Embora reconheça que as empresas beneficiam a sociedade por meio da RSC, Brenkert (1992b) considera um erro as pessoas encorajarem as empresas a se envolverem em RSC como prática. Quando nós fazemos isso, ele diz, nós cedemos uma parte da esfera pública para atores privados. Ao invés de decidirmos juntos como nós desejamos melhorar os males sociais que afetam nossos concidadãos de nossa comunidade, nós deixamos para as organizações privadas decidirem o que fazer. Ao invés de aprimorarmos nossas habilidades democráticas através da deliberação, e reafirmar nossas ligações sociais através de ajuda mútua, nós permitimos que nossas habilidades e vínculos se atrofiem por desuso.

7.2 Empresas, governos e RSC política

Muitas empresas são participantes ativas na arena política. Elas apoiam candidatos para eleição, defendem posições sobre problemas no debate público, fazem lobby junto a autoridades, e mais (vide STARK, 2010). O que a ética dos negócios diz sobre essas atividades?

Cientistas sociais têm produzido uma literatura substancial sobre atividade política corporativa (APC). Para uma revisão, confira o trabalho de Hillman, Keim e Schuler (2004). Essa pesquisa foca questões como: Quais as formas que APC assume? Quais são os antecedentes da APC? Quais são suas consequências? APC levanta muitas questões normativa também.

Uma questão é se empresas são o tipo certo de entidades para se envolver em atividade política. Nos grandes estados, os cidadãos costumam achar útil associar-se outros que pensam da mesma forma que eles, cujo objetivo é representar seus pontos de vista na tomada de decisões políticas. Mas embora organizações como *Republican Party* e a *Sierra Clube* sejam participantes adequados na arena política, não é claro que organizações como Merck ou Wal-Mart o sejam. Isso é porque eles não têm papel reconhecido no processo político, e cidadãos não se juntam neles, nem os deixam com base em considerações políticas (HUSSAIN; MORIARTY, no prelo; TUCKER, 2010). Alguns têm criticado a decisão majoritária da Suprema Corte dos Estados Unidos no *Citizens United*, que afirmou e aprimorou os direitos das empresas de participar de discursos políticos, nessas bases. De modo alternativo, nós deveríamos ver as empresas como oradores legítimos em nome de certos pontos de vista (STARK, 2010).

Especialistas tem também levantando questões sobre os objetivos das APC. Uma coisa que uma empresa deve fazer quando se envolve em APC é fornecer informação valiosa para os agentes do governo. A sociedade tem um interesse em saber como políticas econômicas propostas afetarão as empresas; as próprias empresas são uma boa fonte de informação sobre essas questões. Mas algum temem que empresas se envolvam frequentemente em APC a fim de alcançar seus próprios interesses às custas dos contribuintes. Essa atividade é algumas vezes descrita, e condenada, como rentismo (*rent-seeking*) (TULLOCK, 1989). Questões têm sido levantadas acerca da natureza e permissibilidade do rentismo. De acordo com as definições padrão, o rentismo é uma atividade econômica de desperdício social destinada a garantir benefícios do estado e não do mercado. Mas, há desacordo sobre o que conta como desperdício. Pressões por subsídios agrícolas são frequentemente descritas como uma atividade de busca por rendimentos, mas isso pode ser importante para assegurar o suprimento de comida da nação (BOATRIGHT, 2009b; HINDMOOR, 1999). Uma questão relacionada é se empresas podem se envolver em comportamentos de rentismo. A estrutura do problema aparece ser aquela do dilema do prisioneiro: empresas individuais frequentemente se saem melhor se elas se envolvem no rentismo, mas a economia como um todo se sai pior se todas as empresas se envolverem no rentismo (DEBOW, 1992, 1993).

As formas de APC identificadas acima – participar do discurso público e pressionar oficiais do governo – vai “além” do processo político formal. Mas empresas estão aumentando o engajamento no que parece ser atividade política que vai “no entorno” ou “por fora” desse processo, especialmente em circunstância nas quais o estado é fraco, corrupto ou incompetente. No mundo de hoje, empresas estão provendo bens públicos tais como assistência à saúde e educação (RUGGIE, 2004), protegendo os direitos de cidadania dos indivíduos (MATTEN; CRANE, 2005) e ajudando a criar e a impor sistemas de regulação privados ou “não-vinculativos” (VOGEL, 2010). Por exemplo, quando o Rana Plaza desmoronou em Bangladesch em 2013 matando mais de 1000 trabalhadores da indústria de vestuário, novos códigos de construção e sistemas de aplicação foram postos em prática. Mas eles foram colocados em prática por corporações multinacionais que tem negócios com fábricas de Bangladesh, não pelo governo de Bangladesh. Escritores caracterizam estas atividades como políticas porque elas são do tipo de ações que os estados realizam normalmente, ou deveriam realizar (MATTEN; CRANE, 2005; SCHERER; PALAZZO, 2007, 2011). Estas novas formas de APC – chamadas de “RSC política” – tem levantado questões acerca da legitimidade das ações das empresas em estados governados democraticamente.

Scherer e Palazzo (2007, 2011) são os principais contribuidores para esse debate. Na visão dos mesmos, se uma empresa se comporta como estado, então ela deve ser governada como estado (*vide* MATTEN; CRANE, 2005). A forma de governança que Scherer e Palazzo tem em mente é de caráter habermasiano, envolvendo diálogo deliberativo entre todas as partes interessadas que são afetadas pelas ações da empresa. Eles oferecem como exemplo desse tipo de governança iniciativas *multi-stakeholder* (MSIs) que reúnem empresas, organizações não-governamentais e membros das comunidades locais para deliberar e decidir sobre questões políticas, tais como a *Forest Stewardship Council* (FSC), a *Roundtable on Sustainable Palm Oil* (RSPO), e a *Extractive Industries Transparency Initiative* (EITI). Contra essa visão, críticos têm feito a acusação de que as iniciativas das várias partes interessadas, embora efetivas em produzir um diálogo entre as partes interessadas, são ineficazes em responsabilizar as empresas (MOOG; SPICER; BÖHN, 2015). Há pouquíssimas dúvidas de que empresas podem beneficiar a sociedade através de políticas de RSC. Os códigos de construção estabelecidos pelas multinacionais ocidentais podem muito bem salvar a vida de muitos trabalhadores de vestuário de Bangladesh. A menos que novas formas de governança corporativa possam ser criadas, porém, esses benefícios podem ter um custo para o autogoverno democrático.

Uma maneira ainda mais adequada como as empresas podem se envolver em atividades políticas é através do exercício de seus direitos de propriedade (CHRISTIANO, 2010). Uma empresa pode sair de um estado em resposta à aprovação de uma lei que não é favorável ou pode ameaçar sair de um estado se essa lei for aprovada. Isso pode levar os cidadãos de um estado a revisar ou editar suas decisões políticas. Como em certos casos de políticas de RSC, nós podemos aplaudir o resultado desse tipo de atividade política. Muitos aprovaram quando o estado de Indiana revisou suas leis que permitiam a discriminação contra os membros da comunidade LGBT (sob o fundamento da liberdade religiosa) em resposta a alegações de empresas poderosas, como a *Salesforce.com* e a *Angie's List*, de que reduziram suas atividades econômicas no estado. Mas não é claro se tal comportamento das empresas deveria ser encorajado. Nós podemos desejar estabelecer uma distinção entre indivíduos privados influenciando as tomadas de decisão políticas através do exercício de seus direitos de propriedade e empresas que fazem o mesmo.

7.3 Negócios Internacionais

Muitas empresas transpassam fronteiras sociais – incluindo as nacionais. Esta são chamadas, tipicamente de “multinacionais” ou “transnacionais” (MNCs ou TNCs – sigla em inglês). A operação internacional aumenta a importância de várias questões éticas discutidas acima, como a RSC, mas também levanta novas questões, como o relativismo e o desinvestimento. Duas questões discutidas, frequentemente, em conexão com negócios internacionais não são tratadas nessa seção. Uma é a dos salários e condições de trabalho nas fábricas no exterior, frequentemente chamadas de *sweatshops*. Essa literatura é discutida, brevemente, na seção 6.2. A segunda questão é a corrupção. Para uma discussão sobre essa questão ver a entrada sobre “corrupção”.

Se e em que medida as empresas têm o dever de realizar ações socialmente responsáveis é uma pergunta que pode e foi feita sobre empresas em um contexto doméstico. Mas essa questão parece especialmente urgente em contextos internacionais, e muitos dos exemplos mais famosos de RSC – incluindo o caso da Merck e da Cegueira do Rio discutido na seção 7.1 – acontecem no mundo em desenvolvimento. Há duas razões para isso. Uma é que problemas sociais são, incluindo pobreza e degradação ambiental, piores no mundo em desenvolvimento do que no mundo desenvolvido. A segunda é que as empresas são relativamente atores mais poderosos no mundo em desenvolvimento do que no desenvolvido.

Um número de especialistas em ética dos negócios tem desenvolvido códigos éticos para MNCs, incluindo DeGeorge (1993) e Donaldson (1989). Agências internacionais têm criado códigos de ética para os negócios. Talvez a mais famosa seja a *United Nations Global Compact*, associação na qual se exige que as organizações cumpram uma variedade de regras nas áreas de direitos humanos, trabalho, meio ambiente e combate à corrupção. Em seu relevante trabalho para essa associação, Ruggie (2004, 2013) desenvolveu uma estrutura de “proteção, respeito e reparação” para as MNCs e direitos humanos, que atribui ao Estado o dever primário de proteger e reparar os abusos aos direitos humanos, e às empresas o dever de respeitar os direitos humanos (*vide* WETTSTEIN, 2009). Um fato marcante sobre grande parte dessa pesquisa é que, embora seja focada em negócios internacionais e às vezes promulgada por agências internacionais, a conclusão alcançada não se aplica especificamente a empresas que fazem negócios além das fronteiras nacionais. O dever de, e.g., respeitar direitos humanos aplica-se também a empresas fazendo negócios *dentro* das fronteiras nacionais. É simplesmente

que o contexto internacional é aquele em que esse dever parece ser mais importante a cumprir e em que as empresas são uns dos poucos agentes que podem fazê-lo.

Há problemas, no entanto, que surgem, especialmente, para empresas fazendo negócios internacionalmente. Todo estudante iniciante em ética aprende que diferentes culturas têm diferentes códigos morais. Isso é tipicamente um convite para pensar sobre se a moralidade é ou não relativa à cultura. Para empresários, isso representa um desafio mais imediato: Como deveriam ser gerenciadas as diferenças culturais em códigos morais? Em particular, enquanto operar em um país “anfitrião”, o empresário deveria adotar os padrões desse país, ou deveria aplicar os padrões de seu país “natal”?

Donaldson é uma voz líder sobre essa questão, em trabalhos feitos de modo independente (1998, 1996) e com Dunfee (1999). Donaldson e Dunfee argumentam que há certa “moral mínima” que deve ser seguida em todos os contextos. Ela é dada a nós por “hipernormas”, ou valores e regras morais universais, os quais se justificam por uma “convergência [de sistemas de crenças], religiosas, filosóficas e culturais” (1999, p.57). Dentro das fronteiras do conjunto de Hipernormas, Donaldson e Dunfee dizem que as empresas têm “espaço livre” para selecionar os padrões morais. Elas não têm a liberdade para selecionar qualquer padrão que elas desejam; antes, suas escolhas devem ser guiadas pelas tradições do país anfitrião e de seu nível atual de desenvolvimento econômico.

O enfoque de Donaldson e Dunfee tem atraído uma grande atenção e muitos críticos. Muitas das críticas têm se focado na natureza das hipernormas. Alguns autores argumentam que os critérios de Donaldson e Dunfee para as hipernormas são *ad hoc* (SCHERER, 2015); outros afirmar que algumas das normas que eles afirmam serem hipernormas (por exemplo, a proibição da discriminação de gênero) não satisfazem seus critérios estabelecidos (MAYER; CAVA, 1995). Outros escritores focam a aplicação da teoria de Donaldson e Dunfee, argumentando que ela não dá aos administradores padrões específicos como alegado (SOULE, 2002).

Uma complicação do debate sobre a aplicação dos padrões do país de origem nos países anfitriões é que as empresas multinacionais se envolvem em negócios através das fronteiras nacionais de diferentes maneiras. Algumas MNCs empregam diretamente trabalhadores em vários países, enquanto outras contratam fornecedores em múltiplos países. A Nike, por exemplo, não contrata diretamente trabalhadores para fazer seus tênis. Ao invés disso, a Nike desenha os tênis e contrata empresas em outros países para fazê-los. Nossa opinião sobre se uma multinacional deve aplicar os padrões do país de origem em um país anfitrião pode

depende de se a multinacional os está aplicando a seus próprios trabalhadores ou aos de outras empresas. O mesmo se segue para a responsabilidade. MNCs, especialmente nas indústrias voltadas para o consumidor, muitas vezes são responsabilizadas por más condições de trabalho nas fábricas de seus fornecedores. A Nike foi objeto de críticas afiadas pelas práticas de trabalho de seus fornecedores nos anos 90 (MARTMAN *et al.* 2003). Nossa opinião sobre a extensão da responsabilidade das MNCs pode depender de se as práticas problemáticas existem nas próprias fábricas das MNCs ou nas de seus fornecedores.

Uma empresa pode achar que os padrões de um país anfitrião não são apenas diferentes dos padrões de seu país de origem, mas moralmente intoleráveis. Ela pode decidir que o curso de ação correto é não fazer nenhum negócio no país e, se ela estiver investida no país, desinvestir dele. O problema do desinvestimento recebeu atenção substancial nos anos 80 e 90 com MNCs decidindo se desinvestiam ou não na África do Sul sob o regime do Apartheid. Isso pode atrair atenção renovada nos próximos anos, à medida que empresas e outras organizações contemplam o desinvestimento na indústria de combustíveis fósseis. Razões comuns para desinvestir em uma sociedade ou indústria moralmente problemática são as de evitar cumplicidade nas práticas imorais e pressionar a sociedade ou a indústria a mudar suas práticas. Críticos do desinvestimento preocupam-se acerca dos efeitos dessa prática sobre terceiros partes inocentes (DONALDSON, 1989) e sobre a eficácia do desinvestimento em forçar as mudanças sociais (TEOH; WELCH; WAZZAN, 1999). Alguns acreditam que é melhor para as empresas permanecer envolvidas com a sociedade ou indústria e tentar provocar as mudanças por dentro.

8. O status da ética dos negócios

O campo da ética dos negócios, na sua forma atual, surgiu das pesquisas que filósofos da moral e da política fizeram nos anos 70 e 80. Não é difícil ver por qual motivo filósofos políticos e morais podem estar interessados nos negócios. Atividades comerciais levantam uma série de questões filosóficas interessantes: de agência, confiança, manipulação, exploração, justiça etc.. Após uma onda de atividades há 30 anos, no entanto, os filósofos parecem estar se retirando do campo. Quase não há programas de doutorado em filosofia com professores especializados em ética nos negócios e, como resultado, poucos novos doutores são produzidos nessa área. Aqueles que trabalham na área são, tipicamente, “convertidos” para

teorias éticas e políticas predominantes. Essa é uma oportunidade perdida. Muitos empresários estão preocupados com a ética dos negócios: eles se veem como pessoas boas que querem fazer as coisas certas no trabalho. E muitas agências de credenciamentos, tais como a *Association to Advance Collegiate Schools of Business* (AACSB), exigem que as universidades ensinem ética. Como os filósofos têm abandonado o campo de estudo, universidades focadas em negócios recorrem a especialistas em administração para preencher o vazio. Dado seu treinamento em ciências sociais, professores de administração tratam a ética, em grande parte, como uma empreitada descritiva, ou seja, como um estudo de causas e efeitos dos comportamentos alegadamente ético ou pró-social. Essa é uma importante empreitada, é claro, mas ela não é substituta da reflexão normativa sobre o que é ético nos negócios. Eu espero que esse verbete ajude a informar os filósofos sobre a riqueza e o valor da ética nos negócios, e ao fazer isso, a despertar maior interesse no campo.

Referência bibliográfica

- ALZOLA, M. The Possibility of Virtue. **Business Ethics Quarterly**, v. 22, n. 2, p. 377-404, 2012.
- ANDERSON, E. **Value in Ethics and Economics**. Cambridge: Harvard University Press, 1993.
- ANDERSON, E. **Private Government**: How Employers Rule our Lives (and Why We Don't Talk about It). Princeton: Princeton University Press, 2017.
- ARNESON, R. J. Meaningful Work and Market Socialism. **Ethics**, v. 97, n. 3, p. 517-545, 1987.
- ARNESON, R. J. Meaningful Work and Market Socialism Revisited. **Analyse & Kritik**, v. 31, n. 1, p. 139-151, 2009.
- ARNOLD, D.G.; BOWIE, N. E. Sweatshops and Respect for Persons. **Business Ethics Quarterly**, v. 13, n. 2, p. 221-242, 2003.
- ARNOLD, S. The Difference Principle at Work. **Journal of Political Philosophy**, v. 20, n. 1, p. 94-118, 2012.
- ARRINGTON, R. L. Advertising and Behavior Control. **Journal of Business Ethics**, v. 1, n. 1, p. 3-12, 1982.
- ATTAS, D. What's Wrong with Deceptive Advertising? **Journal of Business Ethics**, v. 21, n. 1, p. 49-59, 1999.

- AYLSWORTH, T. Autonomy and Manipulation: Refining the Argument Against Persuasive Advertising. **Journal of Business Ethics**, 2020. DOI:10.1007/s10551-020-04590-6.
- BAINBRIDGE, S. M. **The New Corporate Governance in Theory and Practice**. New York: Oxford University Press, 2008.
- BARRY, C.; MACDONALD, K. Ethical Consumerism: A Defense of Market Vigilantism. **Philosophy & Public Affairs**, v. 46, n. 3, p. 293-322, 2018.
- BAZERMAN, M. H.; TENBRUNSEL, A. E. **Blind Spots: Why We Fail to Do What's Right and What to Do about It**. Princeton: Princeton University Press, 2011.
- BEBCHUK, L. A.; FRIED, J. M. **Pay Without Performance: The Unfulfilled Promise of Executive Compensation**. Cambridge: Harvard University Press, 2004.
- BERKEY, B. Rawlsian Institutionalism and Business Ethics: Does it Matter Whether Corporations are Part of the Basic Structure of Society? **Business Ethics Quarterly**, v. 31, n. 2, p. 179-209, 2021.
- BLANC, S.; AL-AMOUDI, I. Corporate Institutions in a Weakened Welfare State: A Rawlsian Perspective. **Business Ethics Quarterly**, v. 23, n. 4, p. 497-525, 2003.
- BISHOP, J. D. Is Self-Identity Image Advertising Ethical? **Business Ethics Quarterly**, v. 10, n. 2, p. 371-398, 2000.
- BJÖRNSSON, G.; HESS, K. Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents. **Philosophy and Phenomenological Research**, v. 94, n. 2, p. 273-298, 2017.
- BLAIR, M. M.; STOUT, L. A. A Team Production Theory of Corporate Law. **Virginia Law Review**, v. 85, n. 2, p. 248-320, 1999.
- BOATRIGHT, J. R. Fiduciary Duties and the Shareholder-Management Relation: Or, What's So Special about Shareholders? **Business Ethics Quarterly**, v. 4, n. 4, p. 393-407, 1994.
- BOATRIGHT, J. R. Rent Seeking in a Market with Morality: Solving a Puzzle about Corporate Social Responsibility. **Journal of Business Ethics**, v. 88, n. 4, p. 541-552, 2009b.
- BOATRIGHT, J. R. Executive Compensation: Unjust or Just Right? In: G.G. BRENKERT, G. G.; BEAUCHAMP, T. L. (ed.). **Oxford Handbook of Business Ethics**. New York: Oxford University Press, 2010, p. 161-201.
- BOATRIGHT, J. R.; SMITH, J. D. **Ethics and the Conduct of Business**. 8. ed. Upper Saddle River: Pearson, 2017.

- BOWIE, N. E. **Business Ethics: A Kantian Perspective**. 2. ed. New York: Cambridge University Press, 2017.
- BRATMAN, M. E. Shared Intention. **Ethics**, v. 104, n. 1, p. 97-113, 1993.
- BRENKERT, G. G. Strict Products Liability and Compensatory Justice. *In*: HOFFMAN, W. M.; MOORE, J. M. (ed.). 2. ed. **Business Ethics: Readings and Cases in Corporate Morality**. New York: McGraw-Hill, 1984, p. 460-470.
- BRENKERT, G. G. Private Corporations and Public Welfare. **Public Affairs Quarterly**, v. 6, n. 2, p. 155-168, 1992.
- BRENKERT, G. G. **Marketing Ethics**. Malden: Wiley-Blackwell, 2008.
- BRENKERT, G. G. ISCT, Hypernorms, and Business: A Reinterpretation. **Journal of Business Ethics**, v. 88, n. S4, p. 645-658, 2009.
- BRENKERT, G. G. Whistle-blowing, Moral Integrity, and Organizational Ethics. *In*: BRENKERT, G. G.; BEAUCHAMP, T. L. (ed.). **Oxford Handbook of Business Ethics**. New York: Oxford University Press, 2010, p. 563-601.
- BRENNAN, J.; JAWORSKI, P. M. **Markets Without Limits: Moral Virtues and Commercial Interests**. New York: Routledge, 2016.
- BROWN, B.; MAGUIRE, B. Markets, Interpersonal Practices, and Signal Distortion. **Philosophers' Imprint**, v. 19, n. 4, p. 1-15, 2019.
- CARR, A. Z. Is Business Bluffing Ethical? **Harvard Business Review**, v. 46, n. 1, p. 143-153, 1968.
- CARSON, T. L. **Lying and Deception: Theory and Practice**. New York: Oxford University Press, 2010.
- CHILD, J. W.; MARCOUX, A. M. Freeman and Evan: Stakeholder Theory in the Original Position. **Business Ethics Quarterly**, v. 9, n. 2, p. 207-223, 1999.
- CHRISTIANO, T. The Uneasy Relationship Between Democracy and Capital. **Social Philosophy and Policy**, v. 27, n. 1, p. 195-217, 2010.
- COPP, D. On the Agency of Certain Collective Entities: An Argument for Normative Autonomy. **Midwest Studies in Philosophy**, v. 30, n. 1, p. 194-221, 2006.
- CORVINO, J.; ANDERSON, R. T.; GIRGIS, S. **Debating Religious Liberty and Discrimination**. New York: Oxford University Press, 2017.
- CRISP, R. Persuasive Advertising, Autonomy, and the Creation of Desire. **Journal of Business Ethics**, v. 6, n. 5, p. 413-418, 1987.
- DAHL, R. A. **A Preface to Economic Democracy**. Berkeley: University of California Press, 1985.
- DAVIS, M. Whistleblowing. *In*: LAFOLLETTE, H. (ed.). **Oxford Handbook of Practical Ethics**. New York: Oxford University Press, 2003, p. 539-563.

- DE BRUIN, B. **Ethics and the Global Financial Crisis**. New York: Cambridge University Press, 2015.
- DEGEORGE, R. T. **Competing with Integrity in International Business**. New York: Oxford University Press, 1993.
- DEGEORGE, R. T. **Business Ethic**. 7. ed. Upper Saddle River: Pearson, 2009.
- DELMAS, C. The Ethics of Government Whistleblowing. **Social Theory and Practice**, v. 41, n. 1, p. 77-105, 2015.
- DEMPSEY, J. Corporations and Non-Agential Moral Responsibility. **Journal of Applied Philosophy**, v. 30, n. 4, p. 334-350, 2013.
- DONALDSON, T. **Corporations and Morality**. Englewood Cliffs: Prentice Hall, 1982.
- DONALDSON, T. **The Ethics of International Business**. New York: Oxford University Press, 1989.
- DONALDSON, T. Values in Tension: Ethics Away from Home. **Harvard Business Review**, v. 74, n. 5, p. 48-62, 1996.
- DONALDSON, T. & T.W. DUNFEE, **Ties that Bind: A Social Contracts Approach to Business Ethics**, Cambridge, MA: Harvard Business Press. 1999,
- DONALDSON, T.; WALSH, J. P. Toward a Theory of Business. **Research in Organizational Behavior**, v. 35, p. 181-207, 2015.
- DOW, G. K. **Governing the Firm: Workers' Control in Theory and Practice**. New York: Cambridge University Press, 2003.
- DUSKA, R. Whistleblowing and Employee Loyalty. In: DESJARDINS, J. R.; MCCALL, J. J. (ed.). **Contemporary Issues in Business Ethics**. 4. ed. Belmont: Wadsworth, 2000, p. 167-172.
- DUNFEE, T. W. Do Firms with Unique Competencies for Rescuing Victims of Human Catastrophes Have Special Obligations? Corporate Responsibility and the Aids Catastrophe in Sub-Saharan Africa. **Business Ethics Quarterly**, v. 16, n. 2, p. 185-210, 2006a.
- DUNFEE, T. W. A Critical Perspective of Integrative Social Contracts Theory: Recurring Criticisms and Next Generation Research Topics. **Journal of Business Ethics**, v. 68, n. 3, p. 303-328, 2006b.
- EASTERBROOK, F. H.; FISCHER, D. R. **The Economic Structure of Corporate Law**. Cambridge: Harvard University Press, 1996.
- EDMANS, A.; GABAIX, X.; JENTER, D. Executive Compensation: A Survey of Theory and Evidence. In: HERMALIN, B. E.; WEISBACH, M. S. (ed.). **The Handbook of the Economics of Corporate Governance** (Volume 1). North-Holland: Elsevier, 2017, p. 383-539.

- ELEGIDO, J. M. The Ethics of Price Discrimination. **Business Ethics Quarterly**, v. 21, n. 4, p. 633-660, 2011.
- ELEGIDO, J. M. Does it Make Sense to be a Loyal Employee? **Journal of Business Ethics**, v. 68, n. 3, p. 495-511, 2013.
- EPSTEIN, R. A. In Defense of the Contract at Will. **University of Chicago Law Review**, v. 51, n. 4, p. 947-982, 1984.
- EPSTEIN, R. A. **Forbidden Grounds: The Case Against Employment Discrimination Laws**. Cambridge: Harvard University Press, 1992.
- EVAN, W. M.; FREEMAN, R. E. A Stakeholder Theory of the Modern Corporation: Kantian Capitalism. In: BEAUCHAMP, T. L.; BOWIE, N. E. (ed.). **Ethical Theory and Business**. 3. ed. Englewood Cliffs: Prentice-Hall, 1988, p. 97-106.
- FARACI, D. Wage Exploitation and the Nonworseness Claim. **Business Ethics Quarterly**, v. 29, n. 2, p. 169-188, 2019.
- FERRERAS, I. **Firms as Political Entities: Saving Democracy through Economic Bicameralism**. New York: Cambridge University Press, 2017.
- FREEMAN, R. E. **Strategic Management: A Stakeholder Approach**. Boston: Pitman, 1984.
- FREEMAN, R. E. The Politics of Stakeholder Theory: Some Future Directions. **Business Ethics Quarterly**, v. 4, n. 4, p. 409-421, 1994.
- FREEMAN, R. E.; HARRISON, J. S.; WICKS, A. S.; PARMAR, B. L.; DE COLLE, S. **Stakeholder Theory: The State of the Art**. Cambridge: Cambridge University Press, 2010.
- FREEMAN, R. E.; HARRISON, J. S.; ZYGLIDOPOULOS, S. **Stakeholder Theory: Concepts and Strategies**. New York: Cambridge University Press, 2018.
- FREEMAN, R. E.; REED, D. L. Stockholder and Stakeholders: A New Perspective on Corporate Governance. **California Management Review**, v. 25, n. 3, p. 88-106, 1983.
- FREGA, R.; HERZOG, L.; NEUHÄUSER, C. Workplace Democracy-The Recent Debate. **Philosophy Compass**, v. 14, n. 4, p. 125-174, 2019.
- FRENCH, P. A. The Corporation as a Moral Person. **American Philosophical Quarterly**, v. 16, n. 3, p. 297-317, 1979.
- FRENCH, P. A. **Collective and Corporate Responsibility**. New York: Columbia University Press, 1984.
- FRENCH, P. A. **Corporate Ethics**. Fort Worth: Harcourt Brace, 1995.
- FRIEDMAN, M. The Social Responsibility of Business is to Increase its Profits. **New York Times Magazine**, p. 32-33, p.122-124, 1970.

- GALBRAITH, J. K. **The Affluent Society**. Boston: Houghton Mifflin, 1958.
- GILBERT, M. **Sociality and Responsibility**: New Essays in Plural Subject Theory. Lanham: Rowman & Littlefield, 2000.
- GOLDMAN, A. Ethical Issues in Advertising. *In*: REGAN, T. (ed.). **Just Business**. New York: Random House, 1984, p. 235-270.
- GONZÁLEZ-RICOY, I. The Republican Case for Workplace Democracy. **Social Theory and Practice**, v. 40, n. 2, p. 232-254, 2014.
- GOODPASTER, K. E. Business Ethics and Stakeholder Analysis. **Business Ethics Quarterly**, v. 1, n. 1, p. 53-73, 1991.
- HANSMANN, H. **The Ownership of Enterprise**. Cambridge: Harvard University Press, 1996.
- HANSMANN, H.; KRAAKMAN, R. The End of History for Corporate Law. **Georgetown Law Journal**, v. 89, n. 2, p. 439-468, 2001.
- HARTMAN, E. M. **Virtue in Business**: Conversations with Aristotle. New York: Cambridge University Press, 2015.
- HARTMAN, L. P.; ARNOLD, D. G.; WOKUTCH, R. E. **Rising Above Sweatshops**: Innovative Approaches to Global Labor Challenges. Westport: Praeger, 2003.
- HASAN, R. Rawls on Meaningful Work and Freedom. **Social Theory and Practice**, v. 41, n. 3, p. 477-504, 2015. DOI:10.5840/soctheorpract201541325.
- HASNAS, J. The Normative Theories of Business Ethics: A Guide for the Perplexed. **Business Ethics Quarterly**, v. 8, n. 1, p. 19-42, 1998.
- HASNAS, J. The Mirage of Product Safety. *In*: BRENKERT, G. G.; BEAUCHAMP, T. L. (ed.). **Oxford Handbook of Business Ethics**. New York: Oxford University Press, 2010, p. 677-697.
- HASNAS, J. Reflections on Corporate Moral Responsibility and the Problem Solving Technique of Alexander the Great. **Journal of Business Ethics**, v. 107, n. 2, p. 183-195, 2012.
- HAYEK, F. A. The Use of Knowledge in Society. **American Economic Review**, v. 35, n. 4, p. 519-530, 1945.
- HAYEK, F. A. The Non Sequitur of the Dependence Effect. **Southern Economic Journal**, v. 27, n. 4, p. 346-348, 1961.
- HEATH, J. **Morality, Competition, and the Firm**: The Market Failures Approach to Business Ethics. New York: Oxford University Press, 2014.
- HEATH, J. On the Very idea of a Just Wage. **Erasmus Journal for Philosophy and Economics**, v. 11, n. 2, p. 1-33, 2018.

- HELLMAN, D. **When is Discrimination Wrong?** Cambridge: Harvard University Press, 2008.
- HESS, K. M. The Free Will of Corporations (and Other Collectives). **Philosophical Studies**, v. 168, n. 1, p. 241-260, 2014.
- HINDMOOR, A. Rent Seeking Evaluated. **Journal of Political Philosophy**, v. 7, n. 4, p. 434-452, 1999.
- HOLLEY, D. M. A Moral Evaluation of Sales Practices. **Business & Professional Ethics Journal**, v. 5, n. 1 p. 3-21, 1986.
- HOLLEY, D. M. Information Disclosure in Sales. **Journal of Business Ethics**, v. 17, n. 6, p. 631-641, 1998.
- HSIEH, N.-H. The Obligations of Transnational Corporations: Rawlsian Justice and the Duty of Assistance. **Business Ethics Quarterly**, v. 14, n. 4, p. 643-661, 2004.
- HSIEH, N.-H. Rawlsian Justice and Workplace Republicanism. **Social Theory & Practice**, v. 31, n. 1, p. 115-142, 2005.
- HSIEH, N.-H. Justice in Production. **Journal of Political Philosophy**, v. 16, n. 1, p. 72-100, 2008.
- HSIEH, N.-H.; TOFFEL, M. W.; HULL, O. Global Sourcing at Nike. **Harvard Business School Case Collection**, Case 619-008, 2019.
- HUDSON, R. Ethical Investing: Ethical Investors and Managers. **Business Ethics Quarterly**, v. 15, n. 4, p. 641-657, 2005.
- HUGHES, R. C. Paying People to Risk Life or Limb. **Business Ethics Quarterly**, v. 29, n. 3, p. 295-316, 2019.
- HUGHES, R. C. Pricing Medicine Fairly. **Philosophy of Management**, v. 19, n. 4, p. 369-385, 2020.
- HUSSAIN, W. Corporations, Profit Maximization, and the Personal Sphere. **Economics and Philosophy**, v. 28, n. 3, p. 311-331, 2012.
- HUSSAIN, W. Is Ethical Consumerism an Impermissible Form of Vigilantism? **Philosophy & Public Affairs**, v. 40, n. 2, p. 111-143, 2018.
- HUSSAIN, W.; MORIARTY, J. Accountable to Whom? Rethinking the Role of Corporations in Political CSR. **Journal of Business Ethics**, v. 149, n. 3, p. 519-534, 2018.
- JAWORSKI, P. M. An Absurd Tax on our Fellow Citizens: The Ethics of Rent Seeking in the Market Failures (or Self-Regulation) Approach. **Journal of Business Ethics**, v. 121, n. 3, p. 467-476, 2014.

- JENSEN, M.C. Value Maximization, Stakeholder Theory, and the Corporate Objective Function. **Business Ethics Quarterly**, v. 12, n. 2, p. 235-256, 2002.
- JENSEN, M. C.; MECKLING, M. H. Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure. **Journal of Financial Economics**, v. 3, n. 4, p. 305-360, 1976.
- JONES, T. M.; WICKS, A. C.; FREEMAN, R. E. Stakeholder Theory: The State of the Art. In: BOWIE, N. E. (ed.). **The blackwell guide to business ethics**. Malden: Blackwell, 2002, p. 19-37.
- JONKER, J. The Meaning of a Market and the Meaning of Meaning. **Journal of Ethics and Social Responsibility**, v. 15, n. 2, p. 186-195, 2019.
- KATES, M. The Ethics of Sweatshops and the Limits of Choice. **Business Ethics Quarterly**, v. 25, n. 2, p. 191-212, 2015.
- KOEHN, D.; WILBRATTE, B. A Defense of the Thomistic Concept of the Just Price. **Business Ethics Quarterly**, v. 22, n. 3, p. 501-526, 2012.
- KOLB, R. W. **Too Much is Not Enough**: Incentives in Executive Compensation. New York: Oxford University Press, 2012.
- LANDEMORE, H.; FERRERAS, I. In Defense of Workplace Democracy: Towards a Justification of the Firm-State Analogy. **Political Theory**, v. 44, n. 1, p. 53-81, 2016.
- LAWTON, T.; MCGUIRE, S.; RAJWANI, T. Corporate Political Activity: A Literature Review and Research Agenda. **International Journal of Management Reviews**, 15, n. 1, p. 86-105, 2013.
- LIPPERT-RASMUSSEN, K. **Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination**. New York: Oxford University Press, 2014.
- LIPPKE, R.L. Advertising and the Social Conditions of Autonomy. **Business & Professional Ethics Journal**, v. 8, n. 4, p. 35-58, 1989.
- LIST, C.; PETTIT, P. **Group Agency**: The Possibility, Design, and Status of Corporate Agents. New York: Oxford University Press, 2011.
- MACDONALD, C.; GAVURA, S. Alternative Medicine and the Ethics of Commerce. **Bioethics**, v. 30, n. 2, p. 77-84, 2016.
- MACINTYRE, A.C. **After Virtue**: A Study in Moral Theory. 2. ed. Notre Dame: University of Notre Dame Press, 1984.
- MAITLAND, I. Rights in the Workplace: A Nozickian Argument. **Journal of Business Ethics**, v. 8, n. 12, p. 951-954, 1989.

- MALLESON, T. **After Occupy: Economic Democracy for the 21st Century**. New York: Oxford University Press, 2014.
- MARCOUX, A. M. Much Ado about Price Discrimination. **Journal of Markets & Morality**, v. 9, n. 1, p. 57-69, 2006a.
- MARCOUX, A. M. The Concept of Business in Business Ethics. **Journal of Private Enterprise**, v. 21, n. 2, p. 50-67, 2006b.
- MARX, K. Economic and Philosophical Manuscripts. In: MCLELLAN, D. (ed.). **Karl Marx: Selected Writings**. 2. ed. New York: Oxford University Press, 1844 [2000].
- MASON, A. **Levelling the Playing Field: The Idea of Equal Opportunity and its Place in Egalitarian Thought**. New York: Oxford University Press, 2006.
- MASON, A. Appearance, Discrimination, and Reaction Qualifications. **Journal of Political Philosophy**, v. 25, n. 1, p. 48-71, 2017.
- MAYER, D.; CAVA, A. Social Contract Theory and Gender Discrimination: Some Reflections on the Donaldson/Dunfee model. **Business Ethics Quarterly**, v. 5, n. 2, p. 257-270, 1995.
- MAYER, R. Is There a Moral Right to Workplace Democracy? **Social Theory and Practice**, v. 26, n. 2, p. 301-325, 2000.
- MCCALL, J. J. Employee Voice in Corporate Governance: A Defense of Strong Participation Rights. **Business Ethics Quarterly**, v. 11, n. 1, p. 195-213, 2001.
- MCCALL, J. J.; WERHANE, P. H. Employment at Will and Employee Rights. In: BRENKERT, G. G.; BEAUCHAMP, T. L. (ed.). **Oxford Handbook of Business Ethics**. New York: Oxford University Press, 2010, p. 602-627.
- MCMAHON, C. Morality and the Invisible Hand. **Philosophy and Public Affairs**, v. 10, n. 3 p. 247-277, 1981.
- MCMAHON, C. **Authority and Democracy: A General Theory of Government and Management**. Princeton: Princeton University Press, 1994.
- MCMAHON, C. **Public Capitalism: The Political Authority of Corporate Executives**. Philadelphia: University of Pennsylvania Press, 2013.
- MEJIA, S. Which Duties of Beneficence Should Agents Discharge on Behalf of Principals? A Reflection through Shareholder Primacy. **Business Ethics Quarterly**, p. 1-29, 2020.
- MICHAELSON, C. A Normative Meaning of Meaningful Work. **Journal of Business Ethics**, v. 170, p. 413-428, 2021.
- MILLER, D. **Principles of Social Justice**. Cambridge: Harvard University Press, 1999.

- MILLER, S. Collective Moral Responsibility: An Individualist Account. **Midwest Studies in Philosophy**, v. 30, n. 1, p. 176-193, 2006.
- MOOG, S.; SPICER, A.; BÖHM, S. The Politics of Multi-Stakeholder Initiatives: The Crisis of the Forest Stewardship Council. **Journal of Business Ethics**, v. 128, n. 3, p. 469-493, 2015.
- MOORE, E. S. Children and the Changing World of Advertising. **Journal of Business Ethics**, v. 52, n. 2, p. 161-167, 2004.
- MORIARTY, J. Do CEOs Get Paid Too Much? **Business Ethics Quarterly**, v. 15, n. 2, p. 257-281, 2005a.
- MORIARTY, J. Rawls, Self-Respect, and the Opportunity for Meaningful Work. **Social Theory & Practice**, v. 35, n. 3, p. 441-459, 2009.
- MORIARTY, J. Is 'Equal Pay for Equal Work' Merely a Principle of Nondiscrimination? **Economics and Philosophy**, v. 32, n. 3, p. 435-461, 2016.
- MORIARTY, J. On the Origin, Content, and Relevance of the Market Failures Approach. **Journal of Business Ethics**, v. 165, n. 1, p. 113-124, 2020.
- MULLIGAN, T. **Justice and the Meritocratic State**. New York: Routledge, 2018.
- NORMAN, W. Stakeholder Theory. In: LAFOLLETTE, H. (ed.). **International Encyclopedia of Ethics**. Wiley-Blackwell, 2013.
- NORMAN, W. Rawls on Markets and Corporate Governance. **Business Ethics Quarterly**, v. 25, n. 1, p. 29-64, 2015.
- NOZICK, R. **Anarchy, State, and Utopia**. New York: Basic Books, 1974.
- O'NEILL, M.; WILLIAMSON, T. **Property-Owning Democracy**: Rawls and Beyond. Malden: Wiley-Blackwell, 2012.
- O'NEILL, O. Agents of Justice. **Metaphilosophy**, v. 32, n. 1-2, p. 180-195, 2001.
- ORTS, E. W.; STRUDLER, A. The Ethical and Environmental Limits of Stakeholder Theory. **Business Ethics Quarterly**, v. 12, n. 2, p. 215-233, 2002.
- ORTS, E. W.; STRUDLER, A. Putting a Stake in Stakeholder Theory. **Journal of Business Ethics**, v. 88, n. 4, 605-615, 2009.
- ORTS, E. W.; SMITH, N. C. **The Moral Responsibility of Firms**. New York: Oxford University Press, 2017.
- PAINE, L. S.; BRENKERT, G. G.; WEISSKOFF, R.; KIMMEL, L. D. Children as Consumers: An Ethical Evaluation of Children's Television Advertising [with Commentaries]. **Business & Professional Ethics Journal**, v. 3, n. 3/4, p. 119-169, 1984.

- PALMER, D.; HEDBERG, T. The Ethics of Marketing to Vulnerable Populations. **Journal of Business Ethics**, v. 116, n. 2, p. 403-413, 2013.
- PATEMAN, C. **Participation and Democratic Theory**. New York: Cambridge University Press, 1970.
- PHILLIPS, M. J. The Inconclusive Ethical Case Against Manipulative Advertising. **Business & Professional Ethics Journal**, v.13, n. 4, p. 31-64, 1994.
- PHILLIPS, M. J. Corporate Moral Responsibility: When it Might Matter. **Business Ethics Quarterly**, v. 5, n. 3, p. 555-576, 1995.
- PHILLIPS, R.; FREEMAN, R. E.; WICKS, A. C. What Stakeholder Theory is Not. **Business Ethics Quarterly**, v. 13, n. 4, p. 479-502, 2003.
- POWELL, B.; ZWOLINSKI, M. The Ethical and Economic Case Against Sweatshop Labor: A Critical Assessment. **Journal of Business Ethics**, v. 107, n. 4, p. 449-472, 2012.
- RAWLS, J. **A Theory of Justice**. Cambridge: Harvard University Press, 1971.
- RAWLS, J. **Political Liberalism**. New York: Columbia University Press, 1993.
- ROBSON, G. To Profit Maximize, or Not to Profit Maximize: For Firms, this is a Valid Question. **Economics and Philosophy**, v. 35, n. 2, p. 307-320, 2019.
- ROESSLER, B. Meaningful Work: Arguments from Autonomy. **Journal of Political Philosophy**, v. 20, n. 1, p. 71-93, 2012.
- RÖNNEGARD, D. **The Fallacy of Corporate Moral Agency**. New York: Springer, 2015.
- RUGGIE, J. G. Reconstituting the Global Public Domain: Issues, Actors, and Practices. **European Journal of International Relations**, v. 10, n. 4, p. 499-531, 2004.
- RUGGIE, J. G. **Just Business: Multinational Corporations and Human Rights**. New York: W.W. Norton & Company, 2013.
- SANDEL, M. J. **What Money Can't Buy: The Moral Limits of Markets**. New York: Straus and Giroux, 2012.
- SATZ, D. **Why Some Things Should Not Be For Sale: The Moral Limits of Markets**. New York: Oxford University Press, 2010.
- SCALET, S. P. Fitting the People They Are Meant to Serve: Reasonable Persons in the American Legal System. **Law and Philosophy**, v. 22, n. 1, p. 75-110, 2003.
- SCHARDING, T. K. Imprudence and Immorality: A Kantian Approach to the Ethics of Financial Risk. **Business Ethics Quarterly**, v. 25, n. 2, p. 243-265, 2015.
- SCHERER, A. G. Can Hypernorms be Justified? Insights from a Discourse-Ethical Perspective. **Business Ethics Quarterly**, v. 25, n. 4, p. 489-516, 2015.

- SCHERER, A. G.; PALAZZO, G. The New Political Role of Business in a Globalized World: A Review of a New Perspective on CSR and its Implications for the Firm, Governance, and Democracy. **Journal of Management Studies**, v. 48, n. 4, p. 899-931, 2011.
- SCHOLZ, M.; LOS REYES, G.; SMITH, N. C. The Enduring Potential of Justified Hypernorms. **Business Ethics Quarterly**, v. 29, n. 3, p. 317-342, 2019.
- SCHWARTZ, A. Meaningful Work. **Ethics**, v. 92, n. 4, p. 634-646, 1982.
- SCHWARTZ, D. T. **Consuming Choices**. 2. ed. Lanham: Rowman & Littlefield, 2017.
- SEPINWALL, A. Corporate Moral Responsibility. **Philosophy Compass**, v. 11, n. 1, p. 3-13, 2016.
- SEPINWALL, A. Blame, Emotion, and the Corporation. In: ORTS, E. W.; SMITH, N. C. (ed.). **The Moral Responsibility of Firms**. New York: Oxford University Press, 2017, p. 143-166.
- SHER, S. A Framework for Assessing Immorally Manipulative Marketing Tactics. **Journal of Business Ethics**, v. 102, n. 1, p. 97-118, 2011.
- SILVER, K. Group Action Without Group Minds. **Philosophy and Phenomenological Research**, 2021. DOI:10.1111/phpr.12766.
- SINGER, A. There is No Rawlsian Theory of Corporate Governance. **Business Ethics Quarterly**, v. 25, n. 1, p. 65-92, 2015.
- SINGER, A. **The Form of the Firm: A Normative Political Theory of the Corporation**. New York: Oxford University Press, 2019.
- SISON, A. J. G.; FONTRODONA, J. The Common Good of the Firm in the Aristotelian-Thomistic Tradition. **Business Ethics Quarterly**, v. 22, n. 2, p. 211-246, 2012.
- SMITH, A. **An Inquiry into the Nature and Causes of the Wealth of Nations**. Chicago: University of Chicago Press, 1976.
- SMITH, J.; DUBBINK, W. Understanding the Role of Moral Principles in Business Ethics: A Kantian Perspective. **Business Ethics Quarterly**, v. 21, n. 2, p. 205-231, 2011.
- SNYDER, J. What's the Matter with Price Gouging? **Business Ethics Quarterly**, v. 19, n. 2, p. 275-293, 2009.
- SNYDER, J. Exploitation and Sweatshop Labor: Perspectives and Issue. **Business Ethics Quarterly**, v. 20, n. 2, p. 187-213, 2010.
- SOLOMON, R. C. **Ethics and Excellence: Cooperation and Integrity in Business**. New York: Oxford University Press, 1993.

- STARK, A. Business in Politics: Lobbying and Corporate Campaign Contributions. In: BRENKERT, G. G.; BEAUCHAMP, T. L. (ed.). **Oxford Handbook of Business Ethics**. New York: Oxford University Press, 2010, p. 501-532.
- STEINBERG, E. Big Data and Personalized Pricing. **Business Ethics Quarterly**, v. 30, n. 1, p. 97-117, 2020.
- STERNBERG, E. **Just Business: Business Ethics in Action**. 2. ed. New York: Oxford University Press, 2000.
- STOUT, L. A. **The Shareholder Value Myth: How Putting Shareholders First Harms Investors, Corporations, and the Public**. San Francisco: Berrett-Koehler Publishers, 2012.
- STRUDLER, A. What to Do with Corporate Wealth? **Journal of Political Philosophy**, v. 25, p. 1, p. 108-126, 2017.
- TAYLOR, J. S. **Stakes and Kidneys: Why Markets in Human Body Parts are Morally Imperative**. Burlington: Ashgate Publishing, 2005.
- TULLOCK, G. **The Economics of Special Privilege and Rent Seeking**. Boston: Kluwer Academic, 1989.
- VALDMAN, M. A Theory of Wrongful Exploitation. **The Philosophers' Imprint**, v. 9, n. 6, jul., 2009.
- VARIAN, H. R. Price Discrimination and Social Welfare. **American Economic Review**, v. 75, n. 4, p. 870-875, 1985.
- VELASQUEZ, M. Why Corporations are Not Morally Responsible for Anything They Do. **Business & Professional Ethics Journal**, v. 2, n. 3, p. 1-18, 1983.
- VELASQUEZ, M. Debunking Corporate Moral Responsibility. **Business Ethics Quarterly**, v. 13, n. 4, p. 531-562, 2003.
- VELASQUEZ, M. **Business Ethics: Concepts and Cases**. 7. ed. New York: Pearson, 2012.
- VELTMAN, A. **Meaningful Work**. New York: Oxford University Press, 2016.
- VOGEL, D. **The Market for Virtue: The Potential and Limits of Corporate Social Responsibility**. Washington: Brookings Institution Press, 2005.
- VOGEL, D. The Private Regulation of Global Corporate Conduct: Achievements and Limitations. **Business & Society**, v. 49, n.1, p. 68-87, 2010.
- WARREN, D. E.; GASPAR, J. P.; LAUFER, W. S. Is Formal Ethics Training Merely Cosmetic? A Study of Ethics Training and Ethical Organizational Culture. **Business Ethics Quarterly**, v. 24, n. 1, p. 85-117, 2014.
- WERHANE, P. H. **Persons, Rights, and Corporations**. Englewood Cliffs: Prentice-Hall, 1985.

- WERHANE, P. H.; HARTMAN, L. P.; ARCHER, C. ENGLEHARDT, E. E.; PRITCHARD, M. S. **Obstacles to Ethical Decision-Making: Mental Models, Milgram and the Problem of Obedience**. New York: Cambridge University Press, 2013.
- WETTSTEIN, F. **Multinational Corporations and Global Justice: Human Rights Obligations of a Quasi-Governmental Institution**. Stanford: Stanford Business Books, 2009.
- WOKUTCH, R. E. Nike and its Critics: Beginning a Dialogue. **Organization & Environment**, v. 14, n. 2, p. 207-237, 2001.
- YEOMAN, R. Conceptualising Meaningful Work as a Fundamental Human Need. **Journal of Business Ethics**, v. 125, n. 2, p. 235-251, 2014.
- YOUNG, I. M. **Responsibility for Justice**. New York: Oxford University Press, 2011.
- ZHAO, X.; MURRELL, A. Does a Virtuous Circle Really Exist? Revisiting the Causal Linkage between CSP and CFP. **Journal of Business Ethics**, v. 23, feb., 2021. DOI:10.1007/s10551-021-04769-5.
- ZINGALES, L.; HART, O. Companies Should Maximize Shareholder Welfare not Market Value. **Journal of Law, Finance, and Accounting**, v. 2, n. 2, p. 247-275, 2017.
- ZWOLINSKI, M. Sweatshops, Choice, and Exploitation. **Business Ethics Quarterly**, v. 17, n. 4, p. 689-727, 2007.
- ZWOLINSKI, M. The Ethics of Price Gouging. **Business Ethics Quarterly**, v. 18, n. 3, p. 347-378, 2008.

Sobre os tradutores e revisores

André Nascimento Pontes: Professor da Universidade Federal do Amazonas (UFAM) e do Programa de Pós-Graduação em Filosofia da Universidade Estadual Vale do Acaraú (UVA). Doutor em Filosofia pela Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, com Mestrado e Graduação em Filosofia pela Universidade Federal do Ceará (UFC).

Bruno Aislã Gonçalves dos Santos: Professor da Universidade Estadual do Centro-Oeste (UNICENTRO). Possui Doutorado e Mestrado em Filosofia pela Universidade Federal de Santa Catarina (UFSC). Realizou o doutorado-sanduiche na Universidade de St. Andrews. Estuda o utilitarismo clássico e contemporâneo voltado às discussões sobre a justiça distributiva e a justiça global, dentre outros temas. Atua, principalmente, nas seguintes áreas: Ética, Filosofia Política, Filosofia do Direito, Filosofia da Linguagem, Filosofia da Educação e Filosofia da Economia. Foi bolsista Capes.

Mariana Kuhn de Oliveira: Pesquisadora de Pós-doutorado no Centro Brasileiro de Análise e Planejamento (CEBRAP). Possui Doutorado em Filosofia e em Direito pela Universidade Federal do Rio Grande do Sul (UFRGS) e mestrado em Filosofia pela mesma universidade. Realizou o doutorado-sanduiche na Universidade da Califórnia, em Berkeley. Foi bolsista Capes e Fulbright.

Ramiro de Ávila Peres: Analista de Estudos Especiais do Departamento de Supervisão de Cooperativas e Instituições não-bancárias (Desuc) e Membro do Comitê de Pós-Graduação do Banco Central do Brasil (BCB). Possui graduação em Direito, além de Mestrado e Doutorado em Filosofia, pela Universidade Federal do Rio Grande do Sul (UFRGS). Realizou o doutorado-sanduiche na Universidade Paris I Panthéon-Sorbonne.

Thomas Hyeono Kang: Professor da Escola de Economia de São Paulo, Fundação Getúlio Vargas (FGV EESP) e da Escola Superior de Propaganda e Marketing (ESPM, Porto Alegre). Possui Doutorado em Economia pela Universidade Federal do Rio Grande do Sul (UFRGS) e Mestrado em Economia pela Universidade de São Paulo (USP). Foi *visiting fellow* na School of Advanced Study, da Universidade de Londres. Foi bolsista FAPESP.





DISSERTATIO
FILOSOFIA