

## Democratizing Behavioral Economics

Zachary Liscow & Daniel Markovits<sup>†</sup>

*Behavioral law and economics (“BLE”)—arising from the insight that people make recognizable, systematic mistakes—has revolutionized policymaking. For example, in governments around the world, including the US, teams of experts seek to harness these insights, promising to do things like increase retirement savings. But there is a problem: economic experts do not look or think like the rest of the population. Their demographics and policy views are deeply unrepresentative.*

*This would be less troubling if the experts were merely helping people pursue the behavior that the people themselves would undertake, as was the case in traditional law and economics. However, the whole point of behavioral economics is that such behavior is often not in people’s interest. Rather, in making judgments about the right policy, BLE has erected a new, shaky structure, based on ad hoc and often unstated normative assumptions. The result risks merely enacting the policy preferences (or biases) of unrepresentative experts and thereby distorting policymaking.*

*We propose a new approach—democratic BLE—in which behavioral economists, rather than dictating what the right policy or action is, instead inform representative samples of ordinary people about the evidence, including specifically about their own behavioral biases, and let them decide for themselves. Those decisions, rather than experts’ opinions alone, then inform policymakers. Our approach harnesses the insights of behavioral economics, but in a way that lets the people themselves, rather than the behavioral expert, be the arbiter of the good life.*

---

<sup>†</sup> Yale Law School; zachary.liscow@yale.edu, daniel.markovits@yale.edu. Thanks to Jason Abaluck, Anne Alstott, Oren Bar-Gill, Ryan Bubb, Justin Driver, Lee Fennell, Ray Fisman, Jacob Goldin, Andrew Hayashi, Christine Jolls, Shachar Kariv, Max Kasy, Daryl Levinson, Jeff Rachlinski, Sarath Sanga, Alan Schwartz, Cass Sunstein, Josh Teitelbaum, Kristen Underhill, Glen Weyl, Gui Woolston, and participants in workshops at Boston University Law School, Cornell Law School, ETH Zurich, the University of Hamburg Institute for Law and Economics, and Yale Law School for helpful comments and to Emily Caputo, Audrey Cheng, Mia Dana, David Herman, Eugene Kim, and Sherry Tanious for outstanding research assistance.

Introduction.....	1276
I. Comprehensive Approaches to Measuring Well-Being in BLE.....	1287
A. Structural Estimation .....	1288
B. Modified Revealed Preferences .....	1294
C. Hedonics .....	1297
II. The Medley Approach in Applied BLE: The Case of Retirement	
Savings.....	1302
A. The Continuing Debate on Undersaving for Retirement .....	1303
B. How BLE Claims Undersaving in Practice .....	1307
C. The Cost of Manipulations that Increase Savings .....	1315
D. Other Examples .....	1317
III. Towards a Democratic Behavioral Law and Economics .....	1320
A. Technocratic Behavioral Law and Economics .....	1321
B. Democratic Behavioral Law and Economics .....	1325
C. A Case Study of Democratic BLE .....	1328
D. Democratic BLE in Context.....	1337
Conclusion .....	1338

## Introduction

Economists do not reflect, resemble, or represent the populations that economic policy governs.

On average, economists are more attracted to economic efficiency than ordinary people are.<sup>1</sup> For example, laboratory experiments reveal that economists are less inclined than others to redistribute gains in games from winners to losers.<sup>2</sup> Economists are also more self-interested than other people. In laboratory experiments that offer players a choice about how much of a windfall to keep or share, economists keep more and share less;<sup>3</sup> in experiments that ask players to contribute to a public good, economists free-ride more and contribute less;<sup>4</sup> and in experiments that measure willingness to cheat, economists are more corrupt.<sup>5</sup> This unusual selfishness appears to escape the lab and affect behavior in real life. In one large survey, for example, economists were between twice and eight times as likely to give no money to charity as people educated in other disciplines.<sup>6</sup>

In part, economists' unusual preferences reflect their atypical experiences in becoming and being economists. Educated and high-income elites are generally both more inclined to favor efficiency over equality and more self-interested than other people.<sup>7</sup> And economists necessarily have elite educations and, in our society, capture elite incomes. Salaries for full professors of economics at PhD-granting institutions average over \$215,000,<sup>8</sup> which puts them comfortably in the top 10% of the US income distribution.<sup>9</sup> More narrowly,

---

1. Surveys reveal, for example, that economists are less inclined than others to constrain market prices in the name of fairness. *See, e.g.*, Bruno S. Frey & Werner W. Pommerehne, *On the Fairness of Pricing – An Empirical Survey Among the General Population*, 20 J. ECON. BEHAV. & ORG. 295 (1993); Bruno S. Frey, Werner W. Pommerehne & Beat Gygi, *Economics Indoctrination or Selection? Some Empirical Results*, 24 J. ECON. EDUC. 271 (1993).

2. *See, e.g.*, Reinhard Selten & Axel Ockenfels, *An Experimental Solidarity Game*, 34 J. ECON. BEHAV. & ORG. 517 (1998).

3. *See, e.g.*, John R. Carter & Michael D. Irons, *Are Economists Different, and If So, Why?*, 5 J. ECON. PERSP. 171 (1991).

4. *See, e.g.*, Gerald Marwell & Ruth E. Ames, *Economists Free Ride, Does Anyone Else?*, 15 J. PUB. ECON. 295 (1981).

5. *See, e.g.*, Bjorn Frank & Gunther G. Schulze, *Does Economics Make Citizens Corrupt?*, 43 J. ECON. BEHAV. & ORG. 101 (2000).

6. *See* Robert H. Frank, Thomas Gilovich & Dennis T. Regan, *Does Studying Economics Inhibit Cooperation?*, 7 J. ECON. PERSP. 159 (1993).

7. *See, e.g.*, Raymond Fisman, Pamela Jakiela, Shachar Kariv & Daniel Markovits, *The Distributional Preferences of an Elite*, 349 SCIENCE 1300 (2015); Bryan Caplan, *What Makes People Think Like Economists? Evidence on Economic Cognition from the "Survey of Americans and Economists on the Economy"*, 44 J.L. & ECON. 395, 395 (2001); Bryan Caplan & Stephen C. Miller, *Intelligence Makes People Think Like Economists: Evidence from the General Social Survey*, 38 INTELLIGENCE 636 (2010); *see also* Zachary Liscow & Abigail Pershing, *Why Is So Much Redistribution in Cash and Not In-Kind?*, NAT'L TAX J. (forthcoming 2022) (providing evidence of differing views between the general population and a set of elites on how redistribution should take place).

8. Charles E. Scott & John J. Siegfried, *American Economic Association 2020–2021 Universal Academic Questionnaire Summary Statistics*, 111 AEA PAPERS & PROC. 647, 647–49 (2021).

9. U.S. CENSUS BUREAU, REP. NO. P60-270, HOUSEHOLDS BY TOTAL MONEY INCOME, RACE, AND HISPANIC ORIGIN OF HOUSEHOLDER: 1967 TO 2019 (2020),

economic training in particular emphasizes efficiency-based, and more broadly libertarian, values. As George Stigler long ago observed, the “professional study of economics makes [a person] politically conservative” because it “drill[s]” future economists in “the methods by which a price system solves [economic and political] problems.”<sup>10</sup>

Economists’ unusual commitments also reflect the identities that they bring to their training as economists. A recent survey of economics professors at selective public universities, for example, reports that only 20 percent are female and only 8 percent are Black or Hispanic, while the general US population is 51 percent female and 28.5 percent Black or Hispanic.<sup>11</sup> The pipeline of future economists does not look much more representative: among recent economics PhD recipients, just 34 percent are female and only 11 percent are Black or Hispanic.<sup>12</sup> Economists, that is, are more male and whiter than the population at large—demographic attributes generally correlated with the values that economists also disproportionately embrace.<sup>13</sup> Among US-born PhDs, economics is also the least socioeconomically diverse of major PhD fields, including computer science, math, and other natural and social sciences, in that it has the highest share of PhDs whose parents hold at least a bachelor’s degree (86%) and the highest share whose parents hold at least a graduate degree (65%).<sup>14</sup> These general correlations, moreover, also influence the specific preferences of economists in particular. In the experiments on sharing windfalls and corruption, for example, economists’ unusual reluctance to share and unusual corruptibility were driven by the behavior of male subjects.<sup>15</sup>

A robust and unsettled debate asks whether economists are made through education or instead self-select.<sup>16</sup> But whatever their causes, these differences

---

<https://www.census.gov/data/tables/2020/demo/income-poverty/p60-270.html> [https://perma.cc/MQR6-H87C].

10. George Stigler, *The Politics of Political Economists*, 73 Q.J. ECON. 522, 528 (1959).

11. Diyi Li & Cory Koedel, *Representation and Salary Gaps by Race-Ethnicity and Gender at Selective Public Universities*, 46 EDUC. RESEARCHER 343, 347 tbl.3 (2017).

12. NAT’L SCI. FOUND., NSF 21-308, DOCTORATE RECIPIENTS FROM U.S. UNIVERSITIES tbls.22, 62 (2019).

13. Males and whites are disproportionately conservative. Lydia Saad, *U.S. Conservatism Down Since Start of 2020*, GALLUP (July 27, 2020), <https://news.gallup.com/poll/316094/conservatism-down-start-2020.aspx> [https://perma.cc/S3QT-SENA]. In a pay-what-you-want experiment, men were found to be more selfish than women, resulting in lower payments. Shelle Santana & Vicki G. Morwitz, *The Role of Gender in Pay-What-You-Want Contexts*, 58 J. MARKETING RSCH. 265 (2021). An experiment measuring competitive behavior found men to be more efficiency-minded than women. Loukas Balafoutas, Rudolf Kerschbamer & Matthias Suttera, *Distributional Preference and Competitive Behavior*, 83 J. ECON. BEHAV. & ORG. 125 (2012).

14. Anna Stansbury (@annastansbury), TWITTER (Nov. 9, 2021, 6:05 PM), <https://twitter.com/annastansbury/status/1458209180356055046?> [https://perma.cc/A4EY-XKCW] (discussing a presentation at the Conference on Diversity and Including in Economics, Finance, and Central Banking).

15. See, e.g., Frank & Schulze, *supra* note 5, at 108 tbl.3.

16. See, e.g., Bruno S. Frey & Stephan Meier, *Selfish and Indoctrinated Economists?*, 19 EUR. J.L. & ECON. 165 (2005); Giam Pietro Cipriani, Diego Lubian & Angelo Zago, *Natural Born Economists?*, 30 J. ECON. PSYCHOL. 455 (2009); Justus Haucap & Tobias Just, *Not Guilty? Another Look at the Nature and Nurture of Economics Students* (Deutsche Bank Research Working Paper Series No. 10, 2003), <https://www.econstor.eu/bitstream/10419/40263/1/369887905.pdf> [https://perma.cc/UH67-NQQQ]; Gebhard Kirchgassner, *(Why) Are Economists Different?*, 21 EUR. J. POL. ECON. 543 (2005).

between economists and others—which amount to moral and political disagreements—pose a profound challenge to economics and to the economic analysis of policy in particular. While many economists pursue purely empirical or theoretical questions, economic policy analysis—which we call “law and economics” here—often prescribes how governments should act in practice. Economists (understood broadly to include anyone who deploys economic expertise to analyze policy) do indeed have technical knowledge that others lack. But, if they have unusual moral and political beliefs and preferences, technical knowledge alone is not sufficient to justify their outsized influence<sup>17</sup> over policy that affects everyone. Law and economics therefore faces a profound democratic challenge to its authority.

Law and economics used to have an answer to the challenge of how unrepresentative elites could make policy recommendations:<sup>18</sup> the model of the rational actor.

On the one hand, models of rational choice predict what people will do. A rational actor avoids accidents when liability costs her more than precaution, for example, and she breaches contracts when gains exceed damages. Similarly, in examples to which we shall return below, she changes neighborhoods or cities when new places are sufficiently more appealing than her current situation to outweigh the costs of moving; she studies when the benefits from schooling exceed the burdens of studying; and she saves when the returns from future wealth exceed the sacrifice of foregone present consumption.

On the other hand, the theory of rationality underwrites recommendations about what the law should be. Rational agents are effective at promoting their individual well-being, and welfare economics connects the pursuit of individual welfare to efficient allocations overall.<sup>19</sup> A further set of arguments, concerning freedom and dignity, give these consequentialist ideas deontic support, which many philosophers think provides traditional economics its deepest moral foundations.<sup>20</sup> A rational agent’s freedom entails that her choices deserve deference from the law, which should respect the dignity of autonomous persons. And liberal political theory explains the role that markets play in securing the

17. See, e.g., BINYAMIN APPELBAUM, *THE ECONOMISTS’ HOUR: FALSE PROPHETS, FREE MARKETS, AND THE FRACTURE OF SOCIETY* (2019).

18. See, e.g., David Friedman, *Law and Economics*, in *THE WORLD OF ECONOMICS* 371 (John Eatwell, Murray Milgate & Peter Newman eds., 1991).

19. For example, in the fundamental theorems of welfare economics. See, e.g., Kenneth Arrow, *An Extension of the Basic Theorems of Classical Welfare Economics*, in *PROCEEDINGS OF THE SECOND BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY* 507-32 (Jerzy Neyman ed., 1951); Gerard Debreu, *The Coefficient of Resources Utilization*, 19 *ECONOMETRICA* 273 (1951) (simultaneously offering the first proofs for the modern versions of the welfare theorems); see also KENNETH ARROW & FRANK HAHN, *GENERAL COMPETITIVE ANALYSIS* (1971) (providing the locus classicus for the existence and properties of a competitive equilibrium).

20. See, e.g., JOHN RAWLS, *A THEORY OF JUSTICE* § 2, at 7 (1971) (endorsing market outcomes as long as background justice is maintained through appropriate economic institutions); ARTHUR RIPSTEIN, *FORCE AND FREEDOM: KANT’S LEGAL AND POLITICAL PHILOSOPHY* 232-67 (2009) (defining public goods in legal terms as things that must be provided publicly to ensure the freedom of all members of the political community).

conditions under which the freedom of each is compatible with the equal freedom of all. Insofar as people are rational, therefore, economists can defer to the people themselves by respecting their choices.

These normative traditions face substantial critiques from formidable opponents, of course. As it happens, we find many of the critiques persuasive, and we do not seek here to defend traditional law and economics, all-things-considered. But traditional law and economics is connected to recognizable (and constraining) normative foundations, which reach deep into the history of moral thought. And traditional law and economics has an internal coherence: its descriptive and normative projects work together.

Furthermore, and critically, traditional law and economics makes policy recommendations on grounds that are independent of economists' own particular and often peculiar moral and political preferences. The theory of rationality allows traditional law and economics to deliver concrete outputs—in the form of determinate policy prescriptions—from relatively modest substantive inputs—in the form of assumptions about what is good and how people act. Descriptive work may remain agnostic about what economic actors' interests consist in, secure in the knowledge that they will maximize these interests, whatever they are, and pursue them freely, subject only to the external constraints that they face.<sup>21</sup> Normative work, for its part, need not defend any particular ends or values—it need not decide which accidents should occur and which contracts should be breached, or how much people should study, when they should move, and how much people should save for retirement—because it can rely on rational agents to act authentically and for the best.

The nonrepresentativeness of economists' own moral and political preferences makes this substantive modesty essential for traditional law and economics's normative project. Traditional law and economics scholars can make concrete policy recommendations using parsimonious assumptions *only* because they suppose people to be rational. Rationality allows traditional law and economics to frame itself as the science of helping people to get what *they* want, and this insulates traditional law and economics from attacks that point out that economists' own preferences are unusual. To be sure, traditional law and economics recognizes, and indeed emphasizes, that the rational actor is not *sufficient* for sustaining its normative recommendations. In addition, there must be, for example, perfect competition and no externalities.<sup>22</sup> But those problems leave in place the unified structure: even in their presence, scholars just need to

---

21. Revealed preference theory explains how to discern rational choosers and how to recover their true preferences from their observed choices. S.N. AFRIAT, LOGIC OF CHOICE AND ECONOMIC THEORY (1987) (stating that a finite set of price-quantity observations is consistent with utility maximization if and only if the observations satisfy the Generalized Axiom of Revealed Preference).

22. ANDREU MAS-COLELL, MICHAEL D. WHINSTON & JERRY R. GREEN, MICROECONOMIC THEORY (2012).

observe individuals' rational behavior, thereby learning their preferences and thus the right policy to adopt.<sup>23</sup>

Enter, behavioral economics.

Behavioral economics's first impact is to revolutionize the *description* of human behavior. People are not perfectly rational but instead display known, systematic, and rigorously describable behavioral anomalies and irrationalities.<sup>24</sup> Collectively, these make human action both error-prone and subject to manipulation, including in ways that profoundly influence economic and legal affairs.<sup>25</sup> To predict how people will act, it is not enough to know their true interests and the external constraints that they face, and then to trust that they will rationally optimize these interests given these constraints. Instead, one must also know and understand the behavioral phenomena that, functioning as internal constraints or distortions, cause people to depart from rational optimization. Behavioral economics, alongside allied movements in psychology, sociology, and philosophy, provides the needed understanding. Behavioral law and economics ("BLE") applies this understanding to legal contexts and thus promises to improve on traditional law and economics, as a descriptive project. As Christine Jolls, Cass Sunstein, and Richard Thaler propose in their seminal *A Behavioral Approach to Law and Economics*, "[b]ehavioral economics, in short, offers the potential to be law and economics with a higher 'R<sup>2</sup>'—that is, with greater power to explain observed data."<sup>26</sup>

But BLE remains—prominently—a normative program, which seeks to recommend optimal policies.<sup>27</sup> Normative BLE aspires to identify policies that

23. Bounded rationality is, in this respect, qualitatively different from "imperfect information" or other familiar departures from perfectly functioning markets. Thus, as recent work on information asymmetries between insurance companies and insured parties in health care markets demonstrates, it is perfectly possible to do traditional welfare analysis in a world of imperfect information. Amy Finkelstein, Erzo F.P. Luttmer & Matthew Notowidigdo, *What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption*, 11 J. EUR. ECON. ASS'N 221 (2013).

24. See, e.g., Jennifer Arlen, *The Future of Behavioral Economic Analysis of Law*, 51 VAND. L. REV. 1765, 1766 (1998). Note that a lively debate addresses the relationship between *anomalies* and *irrationalities* and whether particular behavioral departures from the standard model adopted by traditional economics should be regarded as irrational. See, e.g., Kathryn Zeiler, *Mistaken About Mistakes*, 48 EUR. J.L. & ECON. 9 (2019).

25. See, e.g., Cass Sunstein, *Behavioral Analysis of Law*, 64 U. CHI. L. REV. 1175, 1175 (1997).

26. Christine Jolls, Cass Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1487 (1998).

27. See, e.g., B. Douglas Bernheim & Dmitry Taubinsky, *Behavioral Public Economics*, in 1 HANDBOOK OF BEHAVIORAL ECONOMICS: APPLICATIONS AND FOUNDATIONS 381, 434 (B. Douglas Bernheim, Stefano DellaVigna & David Laibson eds., 2018) (describing how to account for bias in making "optimal policy prescriptions"); Raj Chetty, *Behavioral Economics and Public Policy: A Pragmatic Perspective*, 105 AM. ECON. REV. 1, 2 (2015); Sendhil Mullainathan, Joshua Schwartzstein & William Congdon, *A Reduced Form Approach to Behavioral Public Finance*, 4 ANN. REV. ECON. 511, 515-16 (2012) (explaining how behavioral biases play a role in policy-relevant decisions regarding public finance); Richard Thaler & Cass Sunstein, *Behavioral Economics, Public Policy, and Paternalism: Libertarian Paternalism*, 93 AM. ECON. REV. 175, 175 (2003) (noting that "paternalistic" behavioral policies are "selected with the goal of influencing the choices of affected parties in a way that will make those parties better off"); Thomas Ulen, *Behavioral Law and Economics: Law, Policy, and Science*, 21 S. CT. ECON. REV. 5, 9 (2013) (explaining how behavioral economics affects the tools that policymakers should use).

maximize well-being and reflect authentic choices even in the face of the myriad ways in which imperfectly rational agents can, with their behavior, fail to promote their true preferences or to secure their freedom and dignity. BLE-influenced scholarship thus seeks, with increasing success, to influence policy in any number of domains, including consumer protection,<sup>28</sup> public health,<sup>29</sup> crime and policing,<sup>30</sup> and (to conclude with an example that we shall return to repeatedly to develop our arguments) retirement savings.<sup>31</sup> BLE has had particular influence because of the creation of government “nudge units” around the world, including in the US after the promulgation of Executive Order 13,707, directing government agencies to apply the insights of this team of behavioral science experts.<sup>32</sup>

This is natural and proper. Economic policy analysis draws value from its power to drive events and improve outcomes. But the descriptive revolution accomplished by behavioral economics opens a whirlpool beneath the normative foundation of law and economics.<sup>33</sup> Traditional law and economics could avoid becoming entangled in economists’ idiosyncratic moral preferences and beliefs *only* because it assumed that economic agents are perfectly rational, so that their choices reveal their true preferences and therefore identify what their well-being substantively consists in. But as soon as the assumption of rationality is rejected, in the face of BLE’s powerful descriptive results, normative work becomes dependent on a substantive account of well-being. This immediately raises the question: where will this account of well-being come from, and how can it be justified? Given that economists do not reflect, resemble, or represent the general population, the answer cannot be that it comes from the economists’ own intuitions.

The depth of the whirlpool that descriptive BLE opens up beneath normative BLE is not appreciated adequately by many behaviorally influenced law and economics scholars. To be sure, the fact that the behavioral economist can no longer necessarily depend on revealed choices—actual behavior—to identify what substantive outcomes are normatively best has been well known

---

28. See, e.g., OREN BAR-GILL, *SEDUCTION BY CONTRACT: LAW, ECONOMICS, AND PSYCHOLOGY IN CONSUMER MARKETS* (2012).

29. See, e.g., Kristen Underhill, *Extrinsic Incentives, Intrinsic Motivation, and Motivational Crowding-Out in Health Law and Policy*, in *NUDGING HEALTH: HEALTH LAW AND BEHAVIORAL ECONOMICS* (I. Glenn Cohen, Holly Lynch & Christopher Robertson eds., 2016).

30. See, e.g., Richard McAdams & Thomas Ulen, *Behavioral Criminal Law and Economics*, in *CRIMINAL LAW & ECONOMICS* 403-36 (N. Garoupa ed., 2009).

31. See *infra* Section II.B. We choose retirement savings as our central case not because we think it BLE’s most consequential application but rather because so many BLE scholars have addressed the problem, yielding a large reservoir of scholarship and policymaking to engage, and to elaborate critiques that apply generally beyond this case.

32. Exec. Order No. 13,707, 3 C.F.R. § 13707 (2015) (directing federal agencies to employ behavioral insights and this team of experts to “provide agencies with advice and policy guidance to help” do so).

33. We note that even BLE’s descriptive claims remain contested in cases, as for example in recent work calling into question the endowment effect. See, e.g., Kathryn Zeiler, *What Explains Observed Reluctance to Trade? A Comprehensive Literature Review*, in *RESEARCH HANDBOOK ON BEHAVIORAL LAW AND ECONOMICS* 347 (Joshua C. Teitelbaum & Kathryn Zeiler eds., 2018).

from the beginnings of BLE. As Jolls, Sunstein, and Thaler noted early in the movement, behavioral effects “call into question the idea of consumer sovereignty.”<sup>34</sup> But normative BLE has never really come to grips with the consequences of this fact. Instead, well-intentioned behavioral law and economics scholars have come up with a series of back-fillers, which aspire to plug the holes opened up by BLE’s descriptive work and allow BLE scholars to make confident normative recommendations, mostly aimed at saying which laws and policies will maximize well-being.

Two styles of argument dominate these efforts. The first seeks to meet a theoretical challenge with theoretical innovations, by developing new systematic (even comprehensive) approaches to recovering the rationalized preferences of imperfectly rational agents. In Part I, we assess the theoretical techniques that BLE deploys in order to rebuild behavioral law and economics’s normative foundations. We summarize the leading methods by which normative BLE scholars propose to identify the true welfare of imperfectly rational agents. We applaud much of this work but argue that, for all their impressive technical sophistication, none of these methods solves the basic normative problem that BLE faces.

These difficulties entail that, when making policy recommendations in practice, BLE abjures high theory and instead adopts a second style of argument, which we call the “medley approach,” to combine various pieces of evidence to make a policy recommendation.<sup>35</sup> As leading behavioral economist Raj Chetty prescribes, “the decision to include behavioral factors in economic models should be viewed as a pragmatic rather than philosophical choice.”<sup>36</sup> The approach amounts to a kind of multifactor test, which can include various pieces of evidence that are then weighed by the BLE scholar pragmatically. We illustrate the medley approach in Part II, using the behavioral law and economics of retirement savings as a case study.

The medley approach leads to the concern, well-known to lawyers, that multifactor tests allow conclusory cherry-picking that merely rationalizes intuitions behind a smoke-screen of false deliberation. Worse yet, BLE multifactor tests are untethered from even an agreed-upon list of factors; the behavioral scholar can pick and choose not only the weighting of the factors but even the factors themselves. The medley approach therefore requires discounting the revolutionary implications of the descriptive findings of behavioral

---

34. Jolls, Sunstein & Thaler, *supra* note 26, at 1541; *see also* Ryan Bubb & Richard Pildes, *How Behavioral Economics Trims Its Sails and Why*, 127 HARV. L. REV. 1593, 1633 (2014) (“How to measure so-called ‘normative preferences’ reflecting individuals’ true well-being in such contexts is a major challenge, fraught with epistemological difficulties, but there is a growing literature developing a set of methodologies to do so.”). A long list of these epistemological difficulties appears in Mario Rizzo & Douglas Whitman, *The Knowledge Problem of New Paternalism*, 2009 BYU L. REV. 905. Note that some have argued that traditional economics should have narrowly-limited normative ambitions as well. *See, e.g.*, Faruk Gul & Wolfgang Pesendorfer, *Welfare Without Happiness*, 97 AM. ECON. REV. 472 (2007).

35. *See, e.g.*, John Beshears, James J. Choi, David Laibson & Brigitte C. Madrian, *How Are Preferences Revealed?*, 92 J. PUB. ECON. 1787 (2008).

36. Chetty, *supra* note 27, at 28.

economics. Once again, if economists cannot follow the traditional approach and depend upon rational behavior to reveal what promotes well-being, they cannot know reliably what well-being consists in. The medley approach therefore lacks solid foundations. The normative bricolage that the approach imposes turns the BLE scholar, who must choose what evidence to combine in which ways, into a de facto arbiter of the good life, and thus invites paternalism.<sup>37</sup> Worse yet, this paternalism is often grounded in economists' own preferences and moral beliefs, which reflect both the professional deformations that shape opinion among highly educated and highly paid elites in general and economists in particular and the peculiar demographic identities that economists bring to their training and professional lives. The case study of retirement savings illustrates these patterns also.<sup>38</sup>

Our first main contribution, therefore, is to show that the most prominent lines of BLE-influenced normative scholarship developed in response to BLE's descriptive revolution do not in fact meet the challenge that the descriptive revolution poses. Together, Parts I and II demonstrate how the powerful continuing allure of BLE's normative ambitions invites ad hocery and risks BLE-influenced thinkers' inserting personal and class biases into their normative arguments. In this way, we apply some of BLE's descriptive insights to BLE's normative outputs, developing what might be called a behavioral analysis of behavioral law and economics. In so doing, we show that BLE-based policymaking cannot rebuild a firm normative foundation on the cheap and without facing up to the most difficult questions about value. The implications are broad. Most policies—including questions as basic as how many questions to ask on a form, but especially significant questions like savings rates—involve difficult value-tradeoffs, since most policy questions involve resource costs or risk encouraging people to do things that may not be best for them.

Our second main contribution is to propose a new way out of the trap that we expose. Our proposal exploits the fact that BLE's descriptive revolution does not just undermine traditional approaches to normative work in law and economics but also opens up new normative possibilities, which the rational

---

37. If we were starting out BLE today, we would use the gender-neutral term “parentalism” instead of “paternalism,” but we use the standard “paternalism” here.

38. Some BLE scholarship fully retreats from truly normative to merely “prescriptive” argument—in which a goal like “increasing savings” comes, stipulatively, from outside of BLE's methodological frame. This approach constitutes a normative retreat because it demotes economics from its traditional leadership role in policy analysis, leaving the choice of ends to other disciplines and relegating economics to questions concerning means. *See, e.g.*, Jolls, Sunstein & Thaler, *supra* note 26, at 1522 (distinguishing between truly normative argument—which proposes and defends the ends that the law should serve—and “prescriptive” argument—which merely asks “how the law can best be structured to achieve specified ends”). Such analyses are not the target of our analysis, except insofar as they develop understandings of supposed problems to be solved, such as the claim that “[m]any countries are facing a retirement savings crisis.” Shlomo Benartzi & Richard Thaler, *Behavioral Economics and the Retirement Savings Crisis*, 339 *SCIENCE* 1152, 1152 (2013).

actor model previously foreclosed. Where the danger is there grows the saving power also.<sup>39</sup>

The rational actor model made it not just unnecessary but also improper for traditional law and economics to develop a substantive account of well-being. The fact that observed choices reveal rational actors' actual well-being and authentic preferences, thereby making observed choices a reliable guide to policy, also makes other accounts of welfare and freedom that override or constrain revealed preferences unreliable and oppressive. Rational choice theory, in traditional law and economics, is therefore both a license and an imperative: it does not just allow people to choose for themselves but also demands that policy to defer to market choices. This is why traditional law and economics required externalities or other market failures before it would countenance regulation. BLE's descriptive revolution underwrites a converse pair of lessons. By showing that revealed preferences suffer from irrationality and manipulation, descriptive BLE requires a new normative foundation grounded in a substantive account of well-being. Our first contribution is to show that normative BLE, as currently practiced, fails to answer this demand. But at the same time, descriptive BLE, by undermining rational choice theory, also permits other substantive accounts of well-being to guide policy. Our second contribution answers this invitation by developing an alternative way of identifying people's true preferences and promoting their actual well-being. We propose to use BLE's descriptive insights not to nudge or otherwise manipulate ordinary people,<sup>40</sup> but rather to empower them to understand their own biases and to get what they want when they understand themselves. We call our new approach *democratic behavioral law and economics*.

Democratic BLE proposes an agenda for reform—largely internal to the welfarist tradition—through which BLE might combine technical expertise with *citizens' own views* in order to achieve the goals that it has long pursued. Normative BLE's current failings, we argue, stem from its technocratic impulses, which call for experts to deploy BLE's descriptive insights to choose how to manipulate ordinary behavior away from people's revealed preferences and in service of their true preferences. It is the technocratic turn that exposes normative BLE to distortions based on economists' private, undefended, and unusual moral and political preferences. Democratic BLE, by contrast, puts ordinary people in the driver's seat by empowering them to take command of their own choices, even in the face of the behavioral effects that BLE's descriptive agenda exposes. This approach should be incorporated into Executive Order 13,707 governing the "nudge unit" whose expertise administrative agencies are directed to follow.

---

39. MARTIN HEIDEGGER, *THE QUESTION CONCERNING TECHNOLOGY, AND OTHER ESSAYS* 28 (1954) (famously quoting this mantra of Hölderlin).

40. There is a large literature on the ethics of government "manipulation." See generally CASS SUNSTEIN, *THE ETHICS OF INFLUENCE: GOVERNMENT IN THE AGE OF BEHAVIORAL SCIENCE* (2016). Our aim here is not to engage in that important ethical debate. Rather, the Article is primarily concerned with the substantive results of BLE for well-being.

We devote Part III to sketching out this democratic approach, which would complement traditional technocratic BLE in policymaking. Importantly, we do not propose engaging in the short information treatments typical today in behavioral economics. Instead, at its core, democratic BLE puts experts and ordinary people into extensive conversations with each other. Experts engage ordinary people in collective deliberation with a special focus on their own potential behavioral biases; ordinary people express their views on policy with fresh, self-reflective insight; and policymakers make decisions with better knowledge of what individuals in the public, now conscious of their own behavioral biases, actually want for themselves. In this way, democratic BLE aspires to inform policy with well-informed individuals' making choices for themselves, simultaneously self-aware about their own biases and freed from those of the experts.

Democratic BLE is novel but far from unprecedented, in both its ambitions and its methods. The ambition to replace technocratic over-reaching with participatory decision-making is well known in constitutional law, for example. In that area, technocrats who embrace natural law and its secular counterparts propose to determine the scope and content of fundamental constitutional and basic human rights through rational argument by experts. John Rawls, for example, has said that “we submit our conduct to democratic authority only to the extent necessary to share equitably in the inevitable imperfections of a constitutional system.”<sup>41</sup> Popular constitutionalists, by contrast, reject technocracy in favor of democratic self-determination. They reject the narrow limits on participatory politics that Rawls and others propose, and instead make basic rights a matter of democratic sovereign choice.<sup>42</sup> The analogous move to democracy has not been made in law and economics. We make it now.

The methods by which we propose to implement democratic BLE are also familiar and well-tested. In particular, the deliberative polling that we suggest deploying in democratic BLE has worked successfully in a variety of policy areas involving collective political choice, from electricity regulation to budgeting to land use zoning.<sup>43</sup> We adapt these methods to deal specifically with potential *individual behavioral failings* by describing how to engage in individual consciousness-raising and introspection about these failings.<sup>44</sup> We aim, through this innovation, not to demote but to liberate law and economics. By leaving normativity to democratic politics (or its administrative variant), democratic

---

41. RAWLS, *supra* note 20, at 355. Ronald Dworkin would similarly circumscribe the role of the public. RONALD DWORKIN, SOVEREIGN VIRTUE: THE THEORY AND PRACTICE OF EQUALITY 204 (2000).

42. See, e.g., BRUCE ACKERMAN, I WE THE PEOPLE: FOUNDATIONS 6-7 (1991) (discussing dualist democracy); LARRY KRAMER, THE PEOPLE THEMSELVES: POPULAR CONSTITUTIONALISM AND JUDICIAL REVIEW (2004); Robert Post & Reva Siegel, *Popular Constitutionalism, Departmentalism, and Judicial Supremacy*, 92 CALIF. L. REV. 1027 (2004).

43. JAMES FISHKIN, DEMOCRACY WHEN PEOPLE ARE THINKING: REVITALIZING OUR POLITICS THROUGH PUBLIC DELIBERATION 77 (2018).

44. See CENTER FOR DELIBERATIVE DEMOCRACY, <https://cdd.stanford.edu> [<https://perma.cc/Z3JY-LMYC>] (collecting dozens of deliberative polls on collective policy issues but not aimed at increasing self-awareness about potential individual behavioral failings).

BLE lets experts be experts—deploying their training and skill without bearing the burden of legitimating normative tradeoffs that are, ultimately, political. And, given the relatively small cost of engaging in democratic BLE compared to the huge scope of much behaviorally influenced policymaking, the case for making at least some role for it is quite compelling.

Before commencing with the substance of our argument, we conclude this introduction with two observations—one that concerns each of our main contributions.

First, we recognize the peril in criticizing an entire movement, and BLE has in any event been much criticized,<sup>45</sup> including in ways that recognize the challenge of measuring well-being in a behavioral world.<sup>46</sup> We do not claim that every BLE scholar makes every mistake that we identify in every argument that they develop, or even that no BLE scholar ever avoids these mistakes.<sup>47</sup> Rather, we identify the deep problem with how BLE has responded to the sea change that its descriptive findings have unleashed. We show that the descriptive structure of BLE undermines law and economics's inherited normative foundation without developing a general framework capable of building up a new normative foundation in its stead, and that descriptive BLE thereby invites normative ad hocery, which in practice risks simply importing BLE scholars' normative assumptions, unbacked by disciplinary expertise or authority. Certain mistakes therefore filter widely throughout BLE scholarship, including even among scholars who elsewhere say things that acknowledge these mistakes as errors and express an intent to avoid them. The normative errors that we identify are like attractive nuisances or bad habits—almost like behavioral biases. They are difficult to shake reliably, even in those who in some sense know better.

Second, we recognize that it is not possible to build a new intellectual movement out of whole cloth, and certainly not in a single article. We regard our account of democratic BLE as lighting rather than travelling a path. And we hope, by our arguments, to show why it is natural that this path should begin now, in the shadow of BLE's powerful descriptive results, and why it should strive towards a democratic rather than a technocratic goal. A democratic approach to law and economics, when fully worked out, may have a broad scope. But for now, we focus our elaboration of democracy-enhancing techniques specifically on correcting behavioral effects. This connects the new approach to the intellectual developments that make it both necessary and possible.

---

45. See, e.g., Bubb & Pildes, *supra* note 34; Joshua Wright & Douglas Ginsburg, *Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty*, 106 NW. U. L. REV. 1033, 1063 (2012); Richard Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551 (1998).

46. See, e.g., B. Douglas Bernheim & Antonio Rangel, *Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Economics*, 97 Q.J. ECON. 51 (2009).

47. For an example of an exceptionally careful effort, which clearly states its normative assumptions as *assumptions*, see Hunt Allcott, Benjamin B. Lockwood & Dmitry Taubinsky, *Regressive Sin Taxes, With an Application to the Optimal Soda Tax*, 134 Q.J. ECON. 1557 (2019).

## I. Comprehensive Approaches to Measuring Well-Being in BLE

Theorists have proposed three approaches that attempt to give BLE a broad normative foundation. Each aspires to make it possible to measure the well-being of only boundedly rational agents without resorting to ad hoc judgments or normative bricolage.<sup>48</sup>

The first is structural estimation.<sup>49</sup> A structural model seeks to give a particular form of irrationality or nonstandard behavior a mathematically tractable characterization. Insofar as the characterization predicts actual behavior, the BLE analyst then leverages the model to identify the choices that agents would make but for the modelled irrationality and, from this, to discern what outcomes will maximize experienced well-being. By this means, structural estimation aspires to rebuild the inference from observed choices to experienced well-being that the rational actor model provides traditional law and economics and that BLE's descriptive results undermine.

The second approach—modified revealed preference theory—adjusts standard techniques that derive preferences and welfare from choices in order to address imperfectly rational agents.<sup>50</sup> In this case, the BLE analyst identifies rational preferences, whose pursuit maximizes experienced well-being, either by constructing settings in which irrational behaviors do not arise or by identifying agents who remain rational even in settings in which others do not. Then, the analyst analogizes to these constructed settings and agents in order to discern which outcomes maximize experienced well-being for actual agents, whose bounded rationality leads them to choose sub-optimal outcomes in practice.

The third approach—hedonic estimation—ignores agents' choices to measure their experienced well-being directly, by inquiring into their mental states through self-reports of experienced well-being<sup>51</sup> or even directly via chemicals in the brain.<sup>52</sup> BLE then recommends laws and policies that, rather

---

48. We take these three from Chetty, *supra* note 27, at 3, 23-26. Another approach, favored by philosophers, develops a substantive theory of human flourishing that connects an agent's well-being to the possession of certain goods, quite apart from her subjective preferences. *See, e.g.*, DEREK PARFIT, REASONS AND PERSONS (1984); MATTHEW ADLER, WELL-BEING AND FAIR DISTRIBUTION: BEYOND COST-BENEFIT ANALYSIS 159 (2012). But, as Bernheim and Taubinsky note in their only mention of such theories in a 138-page summary of the state of behavioral economics and policy, "[o]bjective theories have received considerably less attention in behavioral public economics than the alternatives." Bernheim & Taubinsky, *supra* note 27, at 385.

49. *See, e.g.*, David Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 Q.J. ECON. 443 (1997).

50. *See, e.g.*, Bernheim & Rangel, *supra* note 46, at 52-53; Jacob Goldin & Daniel Reck, *Revealed-Preference Analysis with Framing Effects*, 128 J. POL. ECON. 2759 (2020); Jacob Goldin, *Which Way to Nudge? Uncovering Preferences in the Behavioral Age*, 125 YALE L.J. 226 (2015) (arguing that nudges should be based on the preferences of nudge-sensitive decisionmakers).

51. *See, e.g.*, Daniel J. Benjamin, Ori Heffetz, Miles S. Kimball & Nichole Szembrot, *Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference*, 104 AM. ECON. REV. 2698 (2014).

52. *See, e.g.*, Daniel Kahneman, *Experienced Utility and Objective Happiness: A Moment-Based Approach*, in CHOICES, VALUES, AND FRAMES 673 (Daniel Kahneman & Amos Tversky eds., 2000).

than tracking or enabling choices, are directly associated with the optimal mental states.

We accept that these theoretical advances yield insights into bounded rationality and that they might help identify which laws and policies best promote the experienced well-being of agents whose choices remain sub-optimal. Nevertheless, these approaches cannot, either severally or taken together, recapture for BLE the normative power and clarity that traditional law and economics asserted before the behavioral revolution. None of the theories supports a general account of how to maximize the well-being or respect the agency of imperfectly rational choosers; and each depends on doubtful and undefended normative assumptions.

#### *A. Structural Estimation*

A structural estimation models an instance of purportedly irrational behavior and then asks how well the model predicts actual choices.<sup>53</sup> The most famous example, developed to model retirement savings, is “hyperbolic discounting.”<sup>54</sup> Standard economic models imagine that individuals—because they would rather consume something now than have to wait to consume it in the future—discount future consumption at a steady rate, known as “exponential discounting.” A person might, for example, value \$1 of consumption next year at 3% less than \$1 of consumption this year, \$1 of consumption the following year at 3% less than that, and so on. The hyperbolic discounting model proposes, instead, that people typically discount the near future, versus the present, at a higher rate and then discount the farther future, versus the near future, at a lower rate.<sup>55</sup> For example, a hyperbolic discounter might value consumption next year at 10% less than consumption this year, but then value consumption in each subsequent year at only 2% less than consumption the year before.

This means that at every moment, people typically wish to save in the future, but put off saving until later. Thus, the agent in the example will not find it worthwhile to save in year 1 in order to consume in year 2, because she discounts year 2 consumption too steeply (by 10%); but she will think, today, that she should save in year 2 in order to consume in year 3, because she discounts that delay barely at all (by just 2%). However, when year 2 actually arrives, she will again put off saving, because she will now steeply discount year 3 consumption (again, by 10%) even as she will then think that she should begin saving in year 3, as she barely discounts the additional delay to year 4 (again, by just 2%). Of course, when year 3 arrives, the dance begins afresh, as it will again, in every subsequent year.

---

53. Chetty, *supra* note 27, at 26-27; Beshears et al., *supra* note 35, at 1790-91.

54. Laibson, *supra* note 49.

55. *See id.* at 449 (noting that events in the near future are discounted at a higher implicit discount rate than events in the distant future).

As David Laibson’s influential presentation of the model vividly imagines (in a formulation we return to presently), people effectively have two selves—the present self (who faces the choice between consumption and saving at the single, current moment), and the future self (who trades off “current” and “future” consumption at many future moments). The present self places an extra discount on the utility of the future self (at every future moment, over and above the discounting that applies across all future moments),<sup>56</sup> with the consequence that although the agent always wants to save in the future, she never saves in the present. As a result, when considering whether to consume today or save for tomorrow, the extra discount on the future causes people to save less than required to maximize the (exponentially) discounted sum of their experienced well-being over time. When the future arrives, therefore, the person experiences regret at having saved less.

Models that deploy hyperbolic discounting, with appropriate parameters, produce predictions that are consistent with actual savings behavior—more consistent, in some cases, than models that treat agents as rational and deploy exponential discounting. Hyperbolic discounting, in other words, is a signal achievement of BLE’s descriptive program—a core instance of law and economics with greater explanatory power.<sup>57</sup> The descriptive success immediately raises a normative puzzle, however. Beshears, Choi, Laibson, and Madrian state the puzzle clearly, when they observe “that non-constant discount rates imply dynamically inconsistent preferences” which lead the agent to “always break[] her previous plans when the moment of action arises.”<sup>58</sup> They propose, therefore, that “revealed preferences *cannot* be a reliable guide to normative preferences,”<sup>59</sup> that is, to the preferences whose satisfaction would promote true well-being.

In this way, hyperbolic discounting produces a conflict between an agent’s present and future selves: either her present consumption will frustrate her future preference for having saved; or forced savings frustrate her present preference for consumption. Critically, the fact that hyperbolic discounting describes and predicts people’s actual savings behavior does not in itself resolve this conflict or settle how much they should optimally save.

Not content with mere description, BLE deploys structural estimations to address this normative question, typically claiming that hyperbolic discounting leads to undersaving. Structural estimation makes it possible to infer from actual choices, rendered mathematically tractable by modelling an “irrationality,” back to the counterfactual choices that agents would have made had they not suffered the estimated “irrationality.” By fitting a model of hyperbolic discounting to actual choices to save for retirement, the BLE analyst can determine by how much a so-called “time-inconsistent” preference for the present over the near

---

56. *Id.* at 443-45.

57. *See* Jolls, Sunstein & Thaler, *supra* note 26, at 1487.

58. Beshears et al., *supra* note 35, at 1790.

59. *Id.* (emphasis added).

future (which produces discounting over and above the “consistent” discounting by which the agent trades off nearer and more distant futures) suppresses savings behavior. This makes it possible to infer how much a person freed of the time-inconsistent preference—who deployed the exponential discount rates that agents use to trade off nearer and more distant futures consistently, including across the present and the near future—would save. In this way, structural estimation aspires to isolate agents’ “irrational” choices from the rest of their choices and to measure their residual “rational” preferences. This allows BLE to use actual savings choices (as in traditional law and economics), even in the face of the inconsistencies that actual agents display, to say just how insufficiently actual hyperbolic discounters in fact save.

This analytic approach to normativity—which erases the initial steeper discount rate from normative analysis and deploys the lower later discount rate right up to the present—privileges the future self over the present self. For example, Laibson, in an article coauthored with Keith Ericson, proposes that the BLE analyst might “simply evaluate welfare from the long-run perspective . . . on the grounds that these are the preferences that are persistent.”<sup>60</sup> Elsewhere, Laibson writes of distortions of time: optical illusion that shrinks or perceptual noise that dampens perceived future benefits.<sup>61</sup> Similarly, Laibson, Andrea Repetto, and Jeremy Tobacman suppose that “[p]eople have a systematic tendency to err . . . in the direction of instantaneous gratification,” so that “few people” claim to be too future-biased: “smoking too few cigarettes, getting to work too early, or watching too little television” and regard such problems as “so unusual that many of them do not even seem intuitively plausible.”<sup>62</sup> Indeed, they even go so far as to associate excessive future-regard with the medically exceptional—with psychiatric disorder, as when they say that “the eating disorder anorexia nervosa may represent one of the few counterexamples” to the general rule that people impatiently pursue present rewards.<sup>63</sup> The widespread use of the normatively-laden term “present bias,” rather than the more neutral and accurate “present focus” to describe the hyperbolic discounter’s mindset shows how commonly hyperbolic discounting is taken to imply irrationality.<sup>64</sup>

However, such reasoning—without more—remains purely stipulative and even conclusory. The reasoning contains no argument to justify which of the agents’ inconsistent preferences are labelled “irrational,” which is why the word appeared in scare quotes earlier. In particular, structural estimations must still

---

60. Keith Ericson & David Laibson, *Intertemporal Choice*, in 2 HANDBOOK OF BEHAVIORAL ECONOMICS, *supra* note 27, at 1, 41.

61. *Id.* at 16.

62. David Laibson, Andrea Repetto & Jeremy Tobacman, *Self-Control and Saving for Retirement*, 1 BROOKINGS PAPERS ON ECON. ACTIVITY 91, 92-93 (1998).

63. *Id.* at 92 n.5. We note that this framing simplifies and likely mischaracterizes the actual psychology of eating disorders.

64. Bernheim & Taubinsky, *supra* note 27, at 392. Bernheim and Taubinsky further note that, “[d]espite widespread use of the phrase ‘present bias’ rather than the more neutral and descriptively accurate ‘present focus,’ the literature offers little in the way of general evidence (not pertaining specifically to addiction) of characterization failure in contemporaneously framed decisions.” *Id.*

explain why a rational savings pattern would “correct for” the agent’s greater preference for the present over all futures, even though doing so frustrates the agent’s current preference for consumption. After all, the estimations could also respect this preference, even when doing so frustrates the agent’s future preferences for having saved, with the result that she will then regret how little she has saved now. Hyperbolic discounters necessarily experience frustration either now or later, and BLE’s normative claims—that hyperbolic discounters save too little—must explain why they should experience frustration now. Mere assertion does not accomplish that.<sup>65</sup> As Douglas Bernheim and Dmitri Taubinsky observe, “[a]s with any economic question, researchers should resolve these issues based on objective, generally applicable criteria informed by pertinent evidence. It (almost) goes without saying that ‘I know it when I see it’ is not a sound methodological principle.”<sup>66</sup>

In their most careful moments, BLE scholars recognize that structural estimation cannot resolve this choice. Bernheim and Taubinsky thus note that “choices in two frames conflict,” and wonder “how can we tell which (if either) accurately reflects preferences, and which is biased?”<sup>67</sup> To be sure, if one *assumes* that a person “has a well-defined unitary objective that reliably guides her choices only when all consequences are delayed, then it is sensible to say that optimization failure can occur when some consequences are immediate.”<sup>68</sup> In this case, the rational self is the future self, and experienced well-being, aggregated over time, is reduced by failing to save more today. “Yet,” Bernheim and Taubinsky observe, “it is also possible” that a person “embraces an objective that guides the choices she makes when actions have immediate consequences.”<sup>69</sup> Indeed, “one could take the position that people achieve true happiness by living in the moment, but that they suffer from a tendency to over-intellectualize when making decisions about the future.”<sup>70</sup> In this case, the steeper discounting that the agent imposes on every future versus the present may capture “her ‘true’ objective,”<sup>71</sup> so that her rational self is her present self.

It is even possible, as the philosopher Harry Frankfurt has observed, that people who defer gratification—including by allowing their future preference for saving tomorrow to lead them, against current impulses, to save today—display weakness of the will, by allowing moralisms spawned of imagined futures to

---

65. Nor does the fact that there are more future days than present ones settle the question. As we argue below, an excessive future-orientation will burden the lived experience of every future date when it becomes the present.

66. Bernheim & Taubinsky, *supra* note 27, at 392. They add that “labeling a model one way rather than another amounts to resolving normative issues by assumption. It is simply too much to hope that choices themselves can reveal which choices are unbiased.” *Id.* at 393.

67. *Id.* at 392.

68. *Id.* at 400.

69. *Id.*

70. *Id.* at 410.

71. *Id.* at 400.

shake the courage of their true and rational current convictions.<sup>72</sup> None other than Adam Smith expressed a similar view, observing in *The Theory of Moral Sentiments* that a person is often:

enchanted with [the] distant idea of felicity. It appears in his fancy like the life of some superior rank of beings, and, in order to arrive at it, he devotes himself for ever to the pursuit of wealth and greatness. . . . And it is well that nature imposes upon us in this manner. It is this deception which rouses and keeps in continual motion the industry of mankind.<sup>73</sup>

Smith was glad that people are deceived by their nature into pursuit of distant wealth and greatness because this deception has positive externalities—it promotes human progress. But for the individual who is deceived, hard work, excessive savings, and deferred gratification in pursuit of future greatness remain irrational—revealed preferences depart from what would maximize experienced well-being over time. In “[r]elabeling the model,” as Bernheim and Taubinsky observe, “one arrives at a different account of true preference and cognitive bias consistent with this alternative perspective.”<sup>74</sup> Critically, “[t]he model itself provides no guidance as to which account is right and which is wrong.”<sup>75</sup> And “[i]n that case,” Bernheim and Taubinsky concede, “references to ‘weakness of will’ reflect disagreements about proper objectives rather than problems with optimization.”<sup>76</sup>

Indeed, it is even possible that neither the current nor the future perspective is exclusively or completely rationalizable, because both accurately reflect the lived experience of a chooser who is not a single, unified, stable agent—at least in the manner required to make aggregating or optimizing welfare across time intelligible. As Bernheim and Taubinsky also admit, the person “may embrace different objectives in different contexts,”<sup>77</sup> so that she has *no* stable, unitary true or rational self. Philosophers have long asked what practical reason requires with respect to an agent’s present concern for her future self. Some have proposed, as one (although far from the only) reasonable answer, that a person should treat tradeoffs between her present and her distant future “selves” as presenting a problem not of rationally optimizing her own interest but rather of how to trade

---

72. See Harry Frankfurt, *Freedom of the Will and the Concept of a Person*, 68 J. PHIL. 5, 11-12 (1971).

73. ADAM SMITH, *THE THEORY OF MORAL SENTIMENTS* part IV, ch. I, 212-14 (Knud Haakonssen ed., Boston University Press 2002) (1790).

74. Bernheim & Taubinsky, *supra* note 27, at 410.

75. *Id.* If people displayed a substantial and clear-cut taste for pre-commitment mechanisms, which force them (perhaps at some future point) to save more than their then-present preferences recommend, then this *might* support the conventional view that associates people’s true selves with their future rather than their present selves. But evidence of the demand for pre-commitment has proved, as Bernheim and Taubinsky say, “elusive.” *Id.* at 463. And even if the evidence were clear, it would not really resolve the normative question, as the taste might itself be a product of excessive and therefore irrational deference to the future, of weakness of the will of the sort Frankfurt observed.

76. *Id.* at 401.

77. *Id.* at 400-01.

off her interests against the interests of another person.<sup>78</sup> This model casts the dual discount rates observed in hyperbolic discounting in a new and very different light. The steeper discount rate imposed between the present and every future measures the self-interested agent's impatience. The flatter discount rate imposed as between nearer and more distant futures measures the other-regarding agent's decreasing altruism with respect to increasingly distant, or strange, other people. "Savings" behavior thus combines a self-interested tradeoff between consuming today rather than tomorrow (true savings), with a pattern of constrained altruism across increasingly distant future "others" (mischaracterized as "savings" by the metaphysical but not normative identity of the saving agent's present and future "selves").

The pattern of behavior associated with hyperbolic discounting, on this interpretation, resembles the pattern of other-regard by which a limited altruist (who cares about other people but less than she cares about herself) strongly prefers herself over all others and then weakly prefers nearer over more distant others (family and friends, for example, over compatriots, and compatriots over strangers). And the fact that future "selves" regret that the present "self" has not saved more reveals not irrationality or the suboptimal pursuit of a unified self-interest but rather the entirely familiar fact that people (the future "selves") wish that others (now the past "self") were more altruistic than they actually are.

Something like this model has also been contemplated by economists, for example when they take Laibson's metaphor of present and future selves normatively seriously. Robert Hall, commenting on Laibson's model, treats savings decisions not as tradeoffs between present and future consumption in a single, optimizing "self" but rather as consumption externalities between distinct persons.<sup>79</sup> In doing so, he follows Laibson's own suggestion that the "savings" problem might be evaluated according to a "multi-self Pareto-criterion, which treats each self as a different person."<sup>80</sup> Of course, this approach dramatically weakens the normative results that BLE can hope to deliver. Notoriously, the Pareto criterion fails to give policy guidance, because it requires that no person is made worse off by a generalizable policy. BLE scholars who embrace this modelling approach therefore effectively throw in the normative towel.<sup>81</sup>

In spite of this, BLE-inspired applied work and especially policy analysis proceeds as if hyperbolic discounting were irrational and normative preferences followed the future self, so that people typically undersave, as we document extensively in our detailed case study of retirement savings below. This

---

78. E.g., PARFIT, *supra* note 48, at 204-09 (understanding connectedness to the future self as the degree to which a person has many strong psychological connections that overlap between the present and future self).

79. Bob Hall, *Self-Control and Saving for Retirement*, 1 BROOKINGS PAPERS ON ECON. ACTIVITY 91 (1998).

80. Ericson & Laibson, *supra* note 60, at 41.

81. One could also ask if there are selves that are preferred among all possible selves. For example, the present self may be rejected by all future selves. However, one of those future selves may also be rejected by both its past and future selves. So, time inconsistency does not easily resolve itself by polling different selves on each other.

widespread but unfounded association of rationality with deferred gratification raises the specter that BLE's policy recommendations simply reprise the biases of those who make them—as instances of what Bernheim and Taubinsky call “type-A paternalism”—that is, “successful workaholics imposing their own personal values on others.”<sup>82</sup> Even this formulation fails to get to the bottom of things, however, as the conception of “success” that it deploys—being tied to measures of wealth and status that implicitly deploy exponentially discounted measures of present value—is itself conclusory with respect to the fundamental question of value at issue.

Structural modelling, because it produces only representation results, cannot advance normative argument beyond such biases. Confident conclusions about optimal savings—and which self, if any, is the true self—require moral arguments about human flourishing over time and (perhaps) fairness across persons. Reasonable minds may differ on such matters. But structural estimation cannot contribute to these debates. There is simply no direct connection between the possibility that behavior can be modeled in this particular way and an evaluation of whether it is good or otherwise desirable.

### *B. Modified Revealed Preferences*

Driven in part by this recognition of the limitations of structural modeling, others have developed techniques for discerning well-being-maximizing behavior from the choices of boundedly rational agents by separating merely expressed from true preferences. These techniques aspire to use empirical methods to identify and to isolate irrationalities—expressed preferences whose satisfaction does not promote experienced welfare—and then to characterize the truly welfare-maximizing preferences that remain once the irrationalities are removed. However, the approaches possess limited normative power, as their inventors themselves acknowledge.

One version of this modified revealed preference approach deploys a frame (i.e., a presentation of information and/or choices) in which revealed preferences are true preferences. We discuss two such examples. The first is by Douglas Bernheim and Antonio Rangel.<sup>83</sup> They stipulate that a certain frame produces unbiased decisions. Having done so, maximizing utility then follows standard rational choice analysis, just taking the preferences revealed in the correct frame as those accurately reflecting welfare. Raj Chetty, Adam Looney, and Kory Kroft apply the methodology to the case of salience in the pricing of goods at supermarkets. They compare purchasing behavior when the tax-inclusive prices are posted to purchasing behavior of the same goods when—as is normally the case in the United States—the posted prices do not reflect sales taxes. They show that consumption goes down when the tax-inclusive prices are salient and

---

82. Bernheim & Taubinsky, *supra* note 27, at 411.

83. Bernheim & Rangel, *supra* note 46.

interpret the behavior under the salient prices as the behavior that maximizes an individual's actual, experienced well-being.<sup>84</sup> This general approach can take a variety of other forms also, including focusing on active decisions, decisions after practice, or aggregating across many people's behavior. In all these varieties, the approach requires deciding which frame reveals truly welfare-maximizing preferences and which distorts expressed preferences so that they depart from experienced well-being. We argue that this introduces a fatal weakness, as there is no non-conclusory way of making the required decision.

A second version of modified revealed preference theory, developed in more recent work by Jacob Goldin and Daniel Reck does not require a frame in which all people are rational.<sup>85</sup> Rather, Goldin and Reck seek empirically to distinguish rational from irrational agents by comparing choices across frames.<sup>86</sup> Specifically, Goldin and Reck propose that people who choose consistently across frames are rational, whereas those whose choices are distorted by mere framing are not.<sup>87</sup> Goldin and Reck then seek to match inconsistent choosers (whose expressed preferences are frame-dependent) with consistent choosers who are similar in other relevant respects.<sup>88</sup> They propose that the true preferences of the inconsistent choosers track the expressed preferences of their consistent analogs and therefore generalize from consistent choosers to maximize total welfare for all.<sup>89</sup> We argue that this approach again requires importing substantive judgments about well-being from outside of economic theory.

Bernheim and Rangel must identify a frame in which people are rational. Such frames may be hard to specify and, in any case, such a frame must be stipulated. The authors' discussion about whether a generalized choice situation is "suspect" is based on the presence of "information processing failures," in which people make mistakes.<sup>90</sup> But if we knew what was a mistake and what was not, then we would not need this method in the first place—indeed, a central lesson of BLE's descriptive agenda is that we cannot discern rational behavior purely empirically, by observing what people do. Accordingly, what constitutes a mistake must be stipulated or somehow inferred. The authors admit as much when they suggest using agreement among "reasonable people" to know what good frames are.<sup>91</sup> In short, determining what a rational frame is requires an assumption or the very kind of guesswork that the authors are trying to avoid. This makes the method conclusory.

---

84. Because non-salient taxes can make taxes less distortionary, despite the over-consumption by individuals, it is not necessarily welfare-maximizing to have fully salient taxes. See Raj Chetty, Adam Looney & Kory Kroft, *Salience and Taxation: Theory and Evidence*, 99 AM. ECON. REV. 1145, 1170-71 (2009).

85. Goldin & Reck, *supra* note 50.

86. *Id.* at 2760.

87. *Id.*

88. *Id.* at 2761.

89. *Id.*

90. Bernheim & Rangel, *supra* note 46, at 85-86.

91. *Id.* at 85.

Take, for example, the case of putting people in a rational frame by increasing the salience of various factors, as Chetty and his coauthors do for the salience of sales taxes. This raises the concern that the very same psychological frame that makes something “salient” has elsewhere been regarded by scholars as leading to *irrational* behavior. Chetty and his coauthors take the salient tax to reveal consumers’ rational preferences—which track experienced well-being.<sup>92</sup> But in other cases, salience has been connected to irrational behavior, in which expressed preferences depart from experienced well-being. Cass Sunstein, for example, suggests that making airplane hijacking salient produces overreactions about terrorism and excessive fear of flying.<sup>93</sup>

Nor is it difficult to imagine an interpretation of Chetty’s results along similar lines. For example, Bernheim and Taubinsky, after observing that “[a]rguably, posting tax-inclusive prices makes the opportunities transparent, while computing them at the register does not,” also observe that tags identifying tax-inclusive pricing “may lead consumers to become especially ‘tax averse,’ for example because the new tags cause them to focus on their resentment of taxes . . . [or] may simply confuse consumers, who might interpret the after-tax prices as before-tax prices, and thus erroneously think the products are more expensive than they actually are.”<sup>94</sup> Yet another interpretation adopts an almost opposite normative perspective. Some people might even prefer to pay money to the government as taxes over paying money to private corporations as prices. As Oliver Wendell Holmes once said, “I like to pay taxes. With them, I buy civilization.”<sup>95</sup> People benefit from the public goods that government provides, support this public provision when they vote for officials who enact taxes, and generally display pro-social preferences, for example, by donating to charity. In the consumer context, tax-exclusive prices frame taxes in ways that emphasize that they go to the state rather than to sellers, whereas tax inclusive prices obscure this distinction in favor of a frame that emphasizes sellers. This raises the possibility that tax-exclusive prices reveal true, rational preferences and that the diminished consumption that Chetty and his coauthors observe in response to tax inclusive pricing might be an irrational response to an obscuring frame. The difficulty that this possibility illustrates is quite general: it is always difficult to know whether a person experiencing salient information is actually acting rationally or not; and deciding this question will generally require input from precisely the sorts of normative judgments that the method purports to provide as outputs.

---

92. Chetty, Looney & Kroft, *supra* note 84.

93. Cass Sunstein, *Terrorism and Probability Neglect*, 26 J. RISK & UNCERTAINTY 121 (2003).

94. See Bernheim & Taubinsky, *supra* note 27, at 440. Bernheim and Taubinsky add that “[o]ne way to justify the paper’s implicit restriction on the welfare-relevant domain would be to show that people are not aware of unposted taxes through surveys. But in fact, the authors demonstrate precisely the opposite using a survey administered to shoppers exiting the store.” *Id.*

95. FELIX FRANKFURTER, MR. JUSTICE HOLMES AND THE SUPREME COURT 71 (1961) (quoting Holmes).

Goldin and Reck’s approach faces an analogous challenge. Specifically, Goldin and Reck require that those who choose consistently across frames are rational. But while this assumption—and Goldin and Reck openly admit that it is just an assumption—might be useful in some settings, it may not be useful in many others, as Goldin and Reck also acknowledge. In particular, people can be consistently *wrong*, both concerning one type of choice across frames or even across many types of choices. People might, for example, consistently undersave or consistently oversave regardless of how the choice between savings and present consumption is presented. Operating across multiple frames, people might buy the wrong home, work the wrong hours, get the wrong amount of education, or consistently make any wrong choice. Consistency across contexts therefore cannot ever yield conclusive proof that the consistent preferences are rational, in the sense of tracking experienced well-being. Goldin and Reck are quite aware of this limitation. Indeed, they end their paper by discussing it, suggesting that the decisions of experts can serve as further benchmarks for rational thought.<sup>96</sup> But this simply concedes the central point—namely that normative work in BLE must get its norms from someplace outside economic thought.

Proponents of the modified revealed preferences approach are largely motivated by avoiding just these discussions, with Bernheim and Rangel, for example, criticizing using “justifications based on loose and inevitably controversial intuition.”<sup>97</sup> They hope, as is sensible in social science, to make progress using uncontroversial, standard tools. We focus, by contrast, on what progress cannot be made in this way. And our focus exposes that modified revealed preferences do not in fact avoid the use of “controversial intuition.” Someone must still stipulate or argue for an assumption of rational behavior in a particular setting. As the authors acknowledge, appeals to “reasonable people” or “experts” may be needed. But that gets them back to the very problem they were trying to solve.

None of this is to say that analyses based on stipulations about what is reasonable provide no useful insights; they do. Nevertheless, modified revealed preference analysis does not answer, or even address, the deepest normative questions so much as pass the buck. As a result, it provides only limited tools for policy guidance.

### C. *Hedonics*

A third approach—hedonics—responds to the difficulty of using observed choices or behavior to infer experienced well-being by trying to measure well-being directly, through simply asking about experience. This approach involves prompting people to report how happy they are, moment-by-moment, or to report their life satisfaction more broadly. Total well-being, according to this view, is

---

96. Goldin & Reck, *supra* note 85, at 2793.

97. Bernheim & Rangel, *supra* note 46, at 51.

simply the sum of individual reported happiness (“felicity functions”), and optimal policy chooses whatever legal rules and regulations maximize the sum.<sup>98</sup> Hedonics harkens back to utilitarian traditions that pre-date the models deployed by traditional law and economics, but it is also gaining influential contemporary adherents. It has been embraced, for example, by economics Nobel Laureate Daniel Kahneman.<sup>99</sup>

The hedonic approach avoids the difficulties associated with individuals who are unable to predict their own well-being by setting aside the study of choices in favor of asking people about experienced well-being and taking them at their word. Based on surveys assessing how people experience their lives, the social planner will then know what maximizes an individual’s well-being and can set policy accordingly. If the results depart from the options that people would have taken if left free to choose, this merely reflects the fact that irrational agents cannot effectively maximize their experienced well-being. In short, a benevolent social planner, armed with hedonic data that reports lived experience, can secure increases in well-being that the agents are blocked by behavioral effects from achieving themselves.

However, the simplicity of this approach covers up a morass of complications, difficulties, and limitations. Most immediately, it is not clear what the surveys at the heart of hedonic estimations evaluate—passing or momentary happiness (pleasure over pain, or some analog), enduring good mood (contentment over frustration), reflective and evaluative life satisfaction (achievement over failure), or something else entirely. Nor is it clear how accurately the surveys evaluate this thing within each individual person or how the surveys create cardinal scales of experience that can sustain the meaningful interpersonal comparisons on which aggregation across persons depends. These challenges are deep and broad. Good climate, for example, seems to improve moment-by-moment happiness<sup>100</sup> even as it does not improve lifetime satisfaction.<sup>101</sup> Contrariwise, having a prestigious job seems to increase lifetime satisfaction, but not moment-by-moment happiness.<sup>102</sup> Hedonic policymaking would have to choose which of these measures to attend to. But the empirical approach cannot possibly resolve the choice—and it does not even try to.

Another challenge to hedonics arises as soon as hedonic prescription is contradicted by choices that people actually make, especially where they sensibly seem to promote some theoretically articulate, or even just intuitively compelling, conception of value. It is difficult to know whether what people do

---

98. Of course, welfares could be weighted.

99. See Kahneman, *supra* note 52. See generally JOHN BRONSTEEN, CHRISTOPHER BUCCAFUSCO & JONATHON S. MASUR, *HAPPINESS AND THE LAW* (2014) (providing a book-length defense of basing the law on happiness).

100. DANIEL KAHNEMAN, *THINKING FAST AND SLOW* 394-95 (2011).

101. *Id.* at 394.

102. Charles Duhigg, *Wealthy, Successful and Miserable*, N.Y. TIMES MAG. (Feb. 21, 2019) <https://www.nytimes.com/interactive/2019/02/21/magazine/elite-professionals-jobs-happiness.html> [<https://perma.cc/G58D-QRZ8>].

and say diverge because their choices are mistaken, or because their experiences remain opaque to them, or because the hedonic measure fails to capture their true goals. This difficulty arises in highly consequential cases. For example, hedonic studies tend to show that having children reduces the happiness of care-giving parents, possibly in ways not compensated for by improvements later in life.<sup>103</sup> But suggestions that having children reduces the well-being, or all-things-considered flourishing, of parents will seem to many people, and with good reason, absurd. Perhaps instead caregivers over-report stress and under-report experienced love, pride, or accomplishment—or parents conceive and raise children with a rational eye towards something besides their own experience of parenting.

Indeed, the traditional theory of rational choice and revealed preferences recognizes and generalizes this possibility when it frames “utility” functions in “as if” terms, saying that rational agents choose “as if” they are maximizing their “utilities,” but expressly acknowledging that these “utilities” do not correspond to any actually experienced mental states, and certainly not to “happiness” in the shallow sense of pleasure over pain. Indeed, and ironically, another prong of behavioral economics, which studies other-regarding preferences, has produced robust results showing that agents engage in rational forms of self-sacrifice, balancing the well-being of others against their own well-being in consistent ways.<sup>104</sup> Nothing in the theory of rationality requires that these choices are only shallowly other-regarding instances of self-serving sacrifice, as if altruism were just a superior way of getting what one wants.<sup>105</sup> For this reason, and indeed for many others, the form of experienced happiness that hedonic studies measure does not exhaust the ends that rational agents pursue.

Another particularly vexing problem for hedonic approaches to measuring well-being concerns adaptive preferences. For example, Amartya Sen has argued that:

[t]he hopeless beggar, the precarious landless labourer, the dominated housewife, [or] the hardened unemployed . . . may all take pleasure in the small mercies, and manage to suppress intense suffering of the necessity of continuing survival, but it would be ethically deeply mistaken to attach a correspondingly small value to the loss of their well-being because of this survival strategy.<sup>106</sup>

That is, the mere fact that people, especially when structurally oppressed, lower their expectations to become satisfied with less does not mean that their reports of experienced happiness correspond to high levels of meaningful well-

---

103. Michael Argyle, *Causes and Correlates of Happiness*, in WELL-BEING: FOUNDATIONS OF HEDONIC PSYCHOLOGY 353, 360 (Daniel Kahneman, Ed Diener & Norbert Schwarz eds., 1999); see also Daniel Kahneman, Alan B. Krueger, David A. Schkade, Norbert Schwarz & Arthur A. Stone, *A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method*, 306 SCIENCE 1776, 1777 (2004) (reporting survey results indicating that taking care of one’s children ranked just above the least enjoyable activities of working, housework, and commuting).

104. See, e.g., Raymond Fisman, Shachar Kariv & Daniel Markovits, *Individual Preferences for Giving*, 97 AM. ECON. REV. 1858 (2007).

105. See PHILIP LARKIN, NONE OF THE BOOKS HAVE TIME (1960), reprinted in THE COMPLETE POEMS OF PHILIP LARKIN 618, 618 (Archie Burnett ed., 2012).

106. AMARTYA SEN, ON ETHICS AND ECONOMICS 45-46 (1987).

being. Oppressed groups who embrace the ideologies that keep them down—for example, women who embrace conceptions of femininity that emphasize service and support to dominant men—may self-report happiness at their own oppression, but that does not mean that they lead flourishing lives.<sup>107</sup> Perhaps less familiarly, economic inequality can also introduce a similar gap between self-reported hedonic states and genuine well-being. Although Sen does not put it this way, the revealed preferences of the billions of poor around the world striving to improve their lot is in dramatic tension with the small declines in happiness that might be stated by those “suppress[ing] intense suffering.”<sup>108</sup> Rather, the well-being of these suffering individuals seems much worse and the revealed preferences of the striving poor around the world suggest that they actually would strongly prefer to be rich.

Moreover, preference adaptation can introduce a rift between reported happiness and genuine well-being even without structural domination and in non-ideological settings. For example, George Loewenstein and Peter Ubel report that there is no empirical difference in long-term well-being of individuals who have had colostomies, a procedure that involves having a bag for human waste permanently attached to one’s body, versus individuals who have not had this procedure.<sup>109</sup> Yet, subjects, when asked how much they would be willing to pay to avoid having a colostomy, report a very large number.<sup>110</sup> Similarly, quadriplegics and paraplegics quickly adapt to their limited capacities and also report no substantial long-term decrease in happiness.<sup>111</sup> Again, there is a tension between stated happiness on the one hand and our intuitions, stated preferences,

107. Bernard Williams describes this pattern of preference in terms of “what may be called *the critical theory principle*, that the acceptance of a justification does not count if the acceptance is itself produced by the coercive power which is supposedly being justified.” BERNARD WILLIAMS, *IN THE BEGINNING WAS THE DEED: REALISM AND MORALISM IN POLITICAL ARGUMENT* 6 (2005). Nancy Fraser applies the sensibilities behind this principle to feminism in particular in NANCY FRASER, *FORTUNES OF FEMINISM: FROM WOMEN’S LIBERATION TO IDENTITY POLITICS TO ANTI-CAPITALISM* (2013).

108. SEN, *supra* note 106, at 45–46.

109. See George Loewenstein & Peter Ubel, *Hedonic Adaptation and the Role of Decision and Experience Utility in Public Policy*, 92 J. PUB. ECON. 1795, 1799 (2008); see also BRONSTEEN ET AL., *supra* note 99, at 17 (explaining that patients with colostomies tend to habituate to their condition, rating their well-being close to those of healthy controls); Norman F. Boyd, Heather J. Sutherland, Karen Z. Heasman, David L. Tritchler & Bernard J. Cummings, *Whose Utilities for Decision Analysis*, 10 J. MED. DECISION MAKING 58, 66 (1990) (finding that patients with colostomies consistently assign higher utility value to life with a colostomy than do healthy patients).

110. See, e.g., Dylan Smith, Ryan Sherriff, Laura Damschroder, George Loewenstein & Peter Ubel, *Misremembering Colostomies? Former Patients Give Lower Utility Ratings Than Do Current Patients*, 25 HEALTH PSYCHOL. 688, 691–93 (2006) (measuring the well-being of people with colostomies against those without and failing to find any significant difference in self-reported mood, even though people with colostomies report that, on average, they would give up almost fifteen percent of their remaining lifespan to regain normal bowel function).

111. MARTIN SELIGMAN, *AUTHENTIC HAPPINESS: USING THE NEW POSITIVE PSYCHOLOGY TO REALIZE YOUR POTENTIAL FOR LASTING FULFILLMENT* 48 (2002) (“Even individuals who become paraplegic as a result of spinal cord accidents quickly begin to adapt to their greatly limited capacities, and within eight weeks they report more net positive emotion than negative emotion. Within a few years, they wind up only slightly less happy on average than individuals who are not paralyzed. Of people with extreme quadriplegia, 84 percent consider their life to be average or above average. These findings fit the idea that we each have a personal set range for our level of positive (and negative) emotion, and this range may represent the inherited aspect of overall happiness.”).

and presumably revealed preferences on the other hand. Any number of carefully articulated philosophical theories of human flourishing suggest that, in these cases at least, preferences rather than stated happiness best track true well-being.

Another example shows how quickly hedonic adaptation can arise and therefore how pervasively reported happiness can mismeasure true well-being, even writ small, in the pleasures and pains of everyday life. Kahneman and coauthors designed an experiment involving “a mild form of torture [that they call] the cold-hand situation.”<sup>112</sup> Each hand of participants in the experiment was exposed to one of two experimental treatments: 1) the hand was placed in unpleasantly cold (14 degree Celsius) water for 60 seconds and then taken out, and the participant was given a warm towel, and 2) the hand was placed in water of the same temperature for the same 60 seconds then an additional thirty seconds as it gradually warmed about one degree, enough to be perceptible to participants. Participants then were told to choose one of the two treatments to endure again.<sup>113</sup> Despite the fact that the overall exposure to cold water was longer in the second treatment, among participants who reported that their pain decreased in the last thirty seconds, fully *eighty percent* chose the second treatment.<sup>114</sup> Of course, this example does not involve a tension strictly between hedonic measures and revealed preferences, so much as a tension between both of these and the strong intuitions that subjects make the “wrong” choice or feel the “wrong” sensation; however, one suspects that subjects would have been very likely to choose the shorter treatment *ex ante*. Indeed, this suspicion is the fulcrum on which the experiment gets its intuitive leverage.

To be sure, cases in which both intuitive and considered judgments of true well-being favor hedonic reporting over revealed preferences also exist. For example, research shows that although people often choose to live far from work, they do not adapt to long commutes but rather are made deeply unhappy by them.<sup>115</sup> We suspect that many people will judge that in this case, true flourishing tracks reported happiness rather than housing choice.<sup>116</sup> Perhaps this judgment reflects a hidden concern for third-party effects, for example concerning the environmental and social costs of sprawl. But even after setting such concerns aside, many observers will doubt that the advantages of living far from work

---

112. KAHNEMAN, *supra* note 100, at 381-82; *see also* Daniel Kahneman, Barbara L. Fredrickson, Charles A. Schreiber & Donald A. Redelmeier, *When More Pain Is Preferred to Less: Adding a Better End*, 4 PSYCHOL. SCI. 401 (1993).

113. *Id.* at 402.

114. *Id.* at 403.

115. Bruno Frey & Alois Stutzer, *Economic Consequences of Mispredicting Utility*, 15 J. HAPPINESS STUD. 937, 953 (2014) (“It is found that full-time workers adapt, to a large extent, to a higher labor income over a period of 3 years. In contrast, people adapt much less to commuting, and even seem to become increasingly sensitive toward the burden of commuting.”).

116. Alois Stutzer & Bruno S. Frey, *Stress That Doesn't Pay: The Commuting Paradox* 20-21 (IZA Discussion Paper No. 1278, 2004) (citations omitted), <https://ftp.iza.org/dp1278.pdf> [<https://perma.cc/8KCM-6AKE>] (“In particular, [people] may make mistakes when they predict their adaptation to daily commuting stress. It has, for example, been found that people do not get used to random noise. In contrast, people adapt to a large extent to higher income. In the case of overestimated adaptation, people systematically choose too long commuting times.”).

truly compensate for the unhappiness brought on by commuting. In any case, the sense that people may be irrationally commuting while they are in fact worse off after having a colostomy is not a matter of sorting better and worse forms of evidence. The colostomy and the commuting evidence both come from hedonics. Rather, in one case, our intuitions do not jibe with the hedonics, while in the other they do.

In short, the hedonics approach in the end is either partial or conclusory.

The approach is partial if judgments of “happiness” are taken at face value, to be of experiences of pleasure over pain only. If the questions used to elicit information for purposes of the hedonic approach carefully and narrowly pick out a present-time mental state like pleasure or joy, then there is no good reason to think that this state should be treated as the dominant good in policy, and there are many reasons to think it should not. Included among these reasons is that reflective and rational agents will say that they freely sacrifice their “happiness” or “joy” for other values, including other experienced mental states such as “satisfaction” or a “sense of accomplishment.” One need not go so far as Nietzsche’s view that “mankind does not strive for happiness; only the Englishman does that”<sup>117</sup> to accept that many other things are valuable and in fact valued, and that an agent who is happy may nevertheless be (and properly regard herself) as badly off in respect of these other values, or vice versa.

On the other hand, the approach is conclusory if “happiness” is taken to mean all-things-considered flourishing. If the questions deployed in hedonic analysis pick out something more complicated, such as flourishing, whose reporting runs through the deliberation and judgment of the agent, then there is no reason to think that the agent’s report of this is any freer of behavioral effects than her choices. Indeed, there is not even a way to decide which perspective hedonic analysis should consult in assessing well-being—whether the authoritative perspective for flourishing is after or before an injury, to name just one example. Finally, any more complex account of flourishing must develop a way to combine flourishing’s many components—pleasure over pain, contentment over frustration, achievement over failure, and perhaps others besides—into an all-things-considered measure. Needless to say, none of these problems can be resolved simply by asking people to report their momentary subjective assessments of well-being.

## II. The Medley Approach in Applied BLE: The Case of Retirement Savings

Behavioral law and economics must pursue its normative projects in the shadow of the limitations of these approaches to measuring well-being in imperfectly rational agents. Typically, normative work that exploits behavioral economics’s descriptive insights responds to this challenge by forsaking theoretical purity and adopting what we will call a “medley approach” to

---

117. FRIEDRICH NIETZSCHE, TWILIGHT OF THE IDOLS 2 (1911).

assessing welfare. That is, BLE “pragmatically” and not “philosophically” combines disparate pieces of evidence from divergent theoretical traditions in a kind of multifactor test, to suss out what maximizes welfare informally without any rationalized or even fully articulate general theory of well-being.

We devote this Part to exploring the medley approach through a close analysis of a case study concerning optimal savings for retirement. One might say that we conduct a behavioral analysis of this instance of behavioral law and economics. Our case study reveals that even as BLE often draws on nothing more than an ad hoc and conclusory assemblage of partial and opportunistic evidence, it presents policy prescriptions as if they reflected a systematic, well-theorized, and fully defended theory of welfare and human flourishing. That is, even as BLE lacks the moral foundation (however contestable) that sustained traditional law and economics, it behaves as if it possesses similarly workable normative foundations. We show how the approach risks turning normative BLE analysis into the imposition of the behavioral analyst’s view of the good life. This risk of normatively biased policy recommendations is especially problematic given how grossly unrepresentative behavioral analysts are, so that their normative views may diverge significantly from those of the general population. We conclude the Part by using examples involving geographic mobility and education to show that this phenomenon is not unique to savings.

#### *A. The Continuing Debate on Undersaving for Retirement*

An intense and long-standing scholarly debate rages about whether workers save enough for retirement, with evidence on both sides. On the one hand, evidence of myopic behavior concerning spending and saving is not hard to come by. One particularly forceful example concerns large drops in consumption at the predictable exhaustion of unemployment benefits.<sup>118</sup> The diminishing marginal utility of consumption entails that people would be better off overall if they smoothed their consumption across the time before and after their benefits expired—and the foreknowledge about precisely when benefits will expire makes this simple to achieve. On the other hand, retirement savings do not arise in a vacuum, as a problem in purely private choice. Instead of being left entirely to their own devices, without any government assistance or intervention, workers face retirement in the context of a wide range of (often mandatory) public programs that—whether by design or as a side-effect—protect the wealth and consumption of the elderly. These range from the Age Discrimination in Employment Act (which functions, partly, as mandatory income protection for older workers),<sup>119</sup> to tax-preferred pensions (which incentivize and increase

---

118. Peter Ganong & Pascal Noel, *Consumer Spending During Unemployment: Positive and Normative Implications*, 109 AM. ECON. REV. 2383 (2019).

119. 29 U.S.C. §§ 621-34 (2018).

retirement savings),<sup>120</sup> to the Social Security program (which disburses approximately one trillion dollars of forced retirement savings every year).<sup>121</sup>

These factors make disagreement about whether people save enough virtually inevitable.

Those who argue that workers do not save enough emphasize that retirement savings are typically not adequate to fund pre-retirement levels of consumption. Some economic models of consumption suggest that optimal savings would fund retirement consumption equal to 100 percent of pre-retirement levels (or possibly even slightly more).<sup>122</sup> However, as one prominent study concludes, many save less than would allow this level of consumption: “[e]ven if households work to age 65 and annuitize all their financial assets, . . . more than half are at risk of being unable to maintain their standard of living in retirement.”<sup>123</sup>

Those who argue that workers do save enough emphasize that poverty rates are considerably higher among the young (aged 18 to 24) than the old (aged over 65)—21 percent versus 9-10 percent—and, indeed, that poverty rates among the elderly are roughly equal to poverty rates among the middle-aged, who are at the peak of their earning powers.<sup>124</sup> They also observe that caloric intake does not drop on retirement, which suggests that retirees do not feel an exceptional financial pinch, at least.<sup>125</sup> Moreover, other careful and systematic models of optimal life-cycle savings and consumption, applied to comprehensive data on

120. *Types of Retirement Plans*, INTERNAL REVENUE SERV. (Aug. 19, 2021), <https://www.irs.gov/retirement-plans/plan-sponsor/types-of-retirement-plans> [https://perma.cc/LP86-R8WR].

121. *Fiscal Year 2020 Budget Overview*, SOC. SEC. ADMIN. (Mar. 2019), [https://www.ssa.gov/budget/FY20Files/2020BO\\_1.pdf](https://www.ssa.gov/budget/FY20Files/2020BO_1.pdf) [https://perma.cc/E5AG-F4TN] (“[W]e will pay . . . a total of about \$950 billion to OASI beneficiaries through the fiscal year.”).

122. See Richard Thaler & Shlomo Benartzi, *Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving*, 112 J. POL. ECON. 164, 182 (2004). See generally Michael Boskin & John Shoven, *Concepts and Measures of Earnings Replacement During Retirement*, in ISSUES IN PENSION ECONOMICS 113 (1987); B. Douglas Bernheim, *Is the Baby Boom Generation Preparing Adequately for Retirement?*, MERRILL LYNCH (1993); Alan Gustman & Thomas Steinmeier, *Effects of Pensions on Saving: Analysis with Data from the Health and Retirement Study* (Nat’l Bureau of Econ. Rsch., Working Paper No. 6681, 1998), <https://www.nber.org/papers/w6681> [https://perma.cc/9938-SND5]. The case that optimal savings would be enough to fund more than 100 percent of pre-retirement consumption rests on survey evidence that households (including those run by economists!) desire an increasing consumption profile over time.

123. Alicia Munnell, Anthony Webb & Francesca Golub-Sass, *The National Retirement Risk Index: An Update*, CTR. FOR RETIREMENT RSCH. AT B.C. 1 (Oct. 2012), [https://crr.bc.edu/wp-content/uploads/2019/12/IB\\_12-20.pdf](https://crr.bc.edu/wp-content/uploads/2019/12/IB_12-20.pdf) [https://perma.cc/M6VK-AL2N].

124. Calculations are based on Census Bureau reporting. See 2018 U.S. CENSUS BUR. AM. CMTY. SURV. B17001, <https://data.census.gov/cedsci/table?tid=ACSDT1Y2018.B17001&vintage=2018> [https://perma.cc/STB7-GP9U]. The poverty rate of 25- to 64-year-olds is between 10% and 12%.

125. See Mark Aguiar & Erik Hurst, *Consumption Versus Expenditure*, 113 J. POL. ECON. 919, 931 (2005). But see Melvin Stephens Jr. & Desmond Toohey, *Changes in Nutrient Intake at Retirement* (Nat’l Bureau of Econ. Rsch., Working Paper No. 24621, 2018). Note that it is not clear what to make of this approach to measuring well-being. Insofar as obesity diminishes welfare, fewer calories might count as higher consumption. And insofar as the rich suffer obesity at lower rates than the rest, this gives the possibility that calories and consumption are inversely correlated empirical support. See Adela Hruby & Frank Hu, *The Epidemiology of Obesity: A Big Picture*, 33 PHARMACOECONOMICS 673, 680 (2016) (explaining that wealth is inversely correlated with obesity).

actual savings, suggest that “84.4 percent of households meet or exceed their wealth targets (and most of those who are below miss by a relatively small amount),” leading modelers to doubt that “the [downward] consumption changes around retirement” that drive claims of suboptimal savings are in fact “due to inadequate accumulation of retirement wealth.”<sup>126</sup>

We do not aspire to resolve or even to litigate this debate. Rather, we note that there exists considerable and reasonable disagreement among economists about whether people’s savings for retirement exhibit any irrationality at all, and even more debate about whether substantial irrational under-saving remains after government intervention.<sup>127</sup> Apart from any ultimate resolution, the existence of this debate in itself raises a profound puzzle for BLE, which we make our focus.

Many BLE analysts—especially government policy teams, but also in universities—nevertheless insist that workers undersave by considerable amounts. Behavioral science teams in governments around the world promote interventions that increase retirement savings to meet an alleged crisis of undersaving.<sup>128</sup> The UK’s Behavioural Insights Team (a government unit testing and recommending behavioral nudges), for example, blithely writes that “[p]sychologists and behavioural economists have consistently found that individuals tend to overweight immediate rewards and heavily discount costs in

126. John Karl Scholz, Ananth Seshadri & Surachai Khittrakun, *Are Americans Saving “Optimally” for Retirement?*, 114 J. POL. ECON. 607, 637-38 (2006); see also Ty Bernicke, *Reality Retirement Planning: A New Paradigm for an Old Science*, 18 J. FIN. PLAN. 56, 56 (2005); Taha Choukhmane, *Default Options and Retirement Saving Dynamics* (unpublished manuscript), <https://tahachoukhmane.com/wp-content/uploads/2019/07/Choukhmane-2019-Default-Options.pdf> [<https://perma.cc/62Y4-H2GK>]; Rowena Crawford & Cormac O’Dea, *Have Households Under-Saved for Retirement?* 17-18 (Netspar Discussion Paper No. 10/2015-061, 2016), <https://ssrn.com/abstract=2743091> [<https://perma.cc/33VH-YJ8Q>] (arguing that among British households approaching pension age, only 7.9 percent are undersaving and the majority are oversaving, with a median surplus of more than 225,000 pounds); Todd Zywicki, *Do Americans Really Save Too Little and Should We Nudge Them to Save More? The Ethics of Nudging Retirement Savings*, 14 GEO. J.L. & PUB. POL’Y 877, 886-89 (2016); Austin Nichols, *Do Financial Planners Advise Us to Save Too Much for Retirement?*, URBAN INST. (Feb. 2012) <https://www.urban.org/sites/default/files/publication/25126/412510-Do-Financial-Planners-Advise-Us-to-Save-Too-Much-for-Retirement-PDF> [<https://perma.cc/WP79-YR7J>]. See generally Greg Kaplan, Giovanni Violante & Justin Weidner, *The Wealthy Hand-to-Mouth*, 2014 BROOKINGS PAPER ON ECON. ACTIVITY 77 (explaining the household finance trio—low liquid wealth, high illiquid wealth, high tendency to consume out of liquid assets—without depending on present-bias or other behavioral effects).

127. John Beshears, James J. Choi, David Laibson & Brigitte C. Madrian, *Behavioral Household Finance*, in 1 HANDBOOK OF BEHAVIORAL ECONOMICS, *supra* note 27, at 177, 179-86.

128. OFF. OF SCI. & TECH. POL’Y, EXEC. OFF. OF THE PRESIDENT, SOCIAL AND BEHAVIORAL SCIENCES TEAM: 2016 ANNUAL REPORT (2016) (leveraging behavioral science research to identify the most effective time to send messages and which program aspects to highlight to encourage enrollment in a federal retirement savings program); Kizzy Gandy, Katy King, Pippa Streeter Hurle, Chloe Bustin & Kate Glazebrook, *Poverty and Decision-Making: How Behavioral Science Can Improve Opportunity in the UK*, BEHAV. INSIGHTS TEAM 43 (OCT. 2016), <http://www.bi.team/wp-content/uploads/2017/02/JRF-poverty-and-decision-making.pdf> [<https://perma.cc/5CE5-8J84>] (concluding that in the areas of credit and savings, the UK government could use behavioral sciences more effectively for policy interventions); Kate Glazebrook, Chris Larkin & Elisabeth Costa, *Improving Engagement with Pension Decisions: The Results from Three Randomised Controlled Trials*, BEHAV. INSIGHTS TEAM 31-32 (Oct. 2017), <https://www.bi.team/wp-content/uploads/2017/10/Pension-wise-trials.pdf> [<https://perma.cc/3TCK-N3V5>] (suggesting providing both an online and phone service, reducing the total amount of information sent to customers, and reducing the use of additional communication channels to spread information in order to promote engagement with the UK government’s Pension Wise service).

the future when making decisions.”<sup>129</sup> The team then continues to observe that “[a] US survey found that individuals who exhibited present bias were more likely to have credit card debt, and have significantly higher levels of credit card debt”<sup>130</sup> and that “individuals who reported preferring to receive lower amounts of money in the present compared to higher amounts in the future were likely to have 19 percent less predicted retirement savings than those that did not display this bias.”<sup>131</sup> The report concludes that “[p]resent bias may explain why individuals in the UK undersave”<sup>132</sup> and, on this basis, recommends manipulations that will increase savings.<sup>133</sup> A US counterpart, appealing to the “widely shared goal” of retiring with financial security, has similarly touted nudges that encouraged Department of Defense servicemembers to increase their retirement savings.<sup>134</sup> Moreover, scholars who turn to giving policy advice become similarly overconfident and unreflective. For example, Nobel Laureate Richard Thaler and coauthor Shlomo Benartzi, writing in *Science*, assert based on thin evidence that “[m]any countries are facing a retirement savings crisis.”<sup>135</sup> And Beshears, Choi, Laibson, and Madrian favor automatic enrollment in retirement savings plans on the ground that “in firms with saving plans that use a non-enrollment (opt-in) default, procrastination may delay enrollment, thereby biasing downward the short-run participation rate.”<sup>136</sup>

These claims of undersavings—at crisis levels—far outstrip the evidence that might support them. Moreover, the inner logics of the policy cases built on top of the assertions are themselves profoundly suspect. For example, the fact, emphasized the UK policy team, that those who report high discount rates also save less merely shows that preferences influence choices, which is completely consistent with rational optimization. And the procrastination cited by Beshears, Choi, Laibson, and Madrian works both ways, so that automatic enrollment manipulates participation rates upward, and their proposed behavioral manipulation will be optimal only if people in fact save irrationally too little, which is precisely the normative question at issue.

This mismatch between the unsettled state of scholarly research about savings adequacy and the strong (almost unanimous) presumption among BLE-inspired policymakers that savings are badly inadequate makes it essential to ask: how exactly do the policymakers know that people do not save enough, and on what basis does BLE justify its program of manipulating behavior to increase savings?

---

129. Gandy et al., *supra* note 128, at 27 (citing work by Laibson, Samuelson, Zeckhauser, Thaler, and others).

130. *Id.* (citing work by Meier & Sprenger).

131. *Id.* at 35 (citing work by Goda, Levy, Manchester & Sojourner).

132. *Id.*

133. *Id.* at 35-38.

134. OFF. OF SCI. & TECH. POL’Y, EXEC. OFF. OF THE PRESIDENT, SOCIAL AND BEHAVIORAL SCIENCES TEAM: 2015 ANNUAL REPORT (2015).

135. Benartzi & Thaler, *supra* note 38, at 1152.

136. Beshears et al., *supra* note 35, at 1791.

*B. How BLE Claims Undersaving in Practice*

To answer this question, we review prominent examples of practical and policy-oriented work that assume people undersave. We do not claim, or even aspire, to take up every consideration addressed by any contribution to this voluminous literature. Instead, we identify and focus on the patterns of argument that dominate leading contributions. We aim to identify a general syndrome that affects a broad field rather than to diagnose specific maladies in particular articles.

We set out by revisiting more closely a leading paper, mentioned earlier, by John Beshears, John Choi, David Laibson, and Brigitte Madrian (“BCLM”), which has been taken to exemplify BLE approaches to retirement savings. In *How Are Preferences Revealed?*,<sup>137</sup> BCLM expressly aspire to measure “normative preferences”—that is, “preferences that represent the economic actor’s true interests”—as distinct from behavior that does not reflect those true interests.<sup>138</sup> They list six “complementary ways of measuring normative preferences,”<sup>139</sup> none of which is ideal but which can be usefully combined. The authors in effect set up a multifactor test, like those frequently used in the law, which identifies forms of inference that might (when artfully combined) sustain a confident assessment of normative preferences, including in order to answer the question whether people in fact undersave for retirement. The six methods are:

*Structural estimation*, as described in the previous Part;

*Active decisions*, or what people do when they must choose and have no default, so that they are not biased by either the status quo or a default and are forced to think about a choice;

*Asymptotic choice*, or what people choose with more experience making a choice, which might reflect greater informedness;<sup>140</sup>

*Aggregated revealed preferences*, or what people in the aggregate choose to do, a kind of “wisdom of the crowds” that resembles a less sophisticated version of the approach of using consistent choosers to help inform what inconsistent choosers want;

*Self-reported preferences*, by which the authors do not mean hedonics (to which some economists pay no heed)<sup>141</sup> but rather statements about what people say they want to do, like quit smoking or save more; and

---

137. See generally *id.*

138. *Id.* at 1787.

139. *Id.* at 1790.

140. This sort of flocking is common. See generally Supreet Kaur, Michael Kremer & Sendhil Mullainathan, *Self-Control at Work*, 123 J. POL. ECON. 1227 (2015).

141. The fact that this six-factor list does not include hedonics begins to give a hint of the arbitrariness involved in the approach.

*Informed preferences*, of two forms: first, those from outside experts like academics or financial planners and, second, decision-makers themselves who are more informed through education or formal training.<sup>142</sup>

We take up the invitation implicit in this elaboration of ways of knowing, by applying the six factors to retirement savings, which is the lead example in BCLM’s article.

1) *Structural estimation*: Structural estimation plays a central role in BLE analysts’ claims that workers save too little. For example, Thaler and Benartzi—in developing “a prescriptive approach” to savings that aspires to “offer advice on how people can improve their decision making and get closer to the normative ideal”<sup>143</sup>—emphasize weakness of the will, procrastination, and hyperbolic discounting.<sup>144</sup> But these approaches, for all their descriptive power, cannot provide the undersaving hypothesis with stable normative foundations. As described in the previous Part, structural estimations in general require additional, outside premises in order to distinguish which behaviors reflect normative preferences and which reflect behavioral irrationalities. In addition, retirement savings in particular has produced dueling structural estimations—as some affirm and others reject undersaving.<sup>145</sup> The first factor therefore leaves the normative question whether workers in fact undersave in *equipoise*.

2) *Active decisions*: Behavioral economists have documented that the defaults set by employer-provided savings plans powerfully influence workers’ savings behavior. Often, BLE analysts treat this fact as a demonstration that workers who face (conventional) opt-in plans save too little and that opt-out plans would induce more nearly “optimal” savings. As Benartzi and Thaler put the point, “[t]here is now conclusive evidence that automatic enrollment, where employees are automatically signed up unless they opt out, is extremely successful in overcoming the procrastination that can impede signing up.”<sup>146</sup> The inference behind this reasoning is simply invalid, however. When they write in this way, Benartzi and Thaler associate “procrastination” asymmetrically with irrational failures to opt *into* savings; but in fact procrastination applies just as powerfully to produce irrational failures to opt *out*, as BLE studies documenting the stickiness of opt-in defaults demonstrate. To be sure, workers who are defaulted into a savings plan save more than workers who are defaulted out of a plan. But the normative question then becomes whether the bigger problem is that people save too much when defaulted into a plan or that people save too little when defaulted out. Seeing whether the results of active decisions are closer to

---

142. Beshears et al., *supra* note 35, at 1790-93.

143. See Thaler & Benartzi, *supra* note 122, at 167.

144. *Id.* at 166-70.

145. Compare Alicia Munnell, Anthony Webb & Francesca Golub-Sass, *Is There Really a Retirement Savings Crisis? An NRRRI Analysis*, CTR. FOR RETIREMENT RSCH. AT B.C. 1 (Aug. 2007), [https://crr.bc.edu/wp-content/uploads/2007/07/ib\\_7-11-508.pdf](https://crr.bc.edu/wp-content/uploads/2007/07/ib_7-11-508.pdf) [<https://perma.cc/N2UA-UBT3>] (finding an increasing number of Americans are undersaving for retirement), with Scholz, Seshadri & Khitatrakun, *supra* note 126, at 637 (finding little evidence that households nearing retirement undersave).

146. Benartzi & Thaler, *supra* note 38, at 1152.

the opt-in or opt-out savings rates can provide insight to this question. And BCLM report that, in practice, active savings decisions produce behavior that is about midway between the observed behavior under opt-in and opt-out regimes.<sup>147</sup> Overall then, neither under- nor oversaving seems a clearly greater problem, and the normative question remains in *equipoise*.

3) *Asymptotic choice*: This factor considers how experience leads behavior to change over time. In applying this factor, it is essential to fix a frame in which to measure experience. In one sense, people have a great deal of (life-long) experience at saving, and this leads them to save as they actually do. Insofar as experience tutors reason, then, the stability of actual savings behavior cuts against any suggestion that people irrationally save too little. Narrower time-frames do not conclusively change this calculus. On the one hand, nudges that increase savings seem to do so stably over at least the middle term. In one study in which workers pre-committed to devoting future raises to retirement savings, for example, a large “majority of these participants did not change their mind once the savings increases took place.”<sup>148</sup> On the other hand, another study of the effect of auto-enrollment in retirement savings plans concludes that, over time, though people save more in specifically designated “retirement accounts,” these increases in savings are fully offset by increased borrowing for things like cars.<sup>149</sup> Thus, even if workers do not leave autoenrollment over time, this does not necessarily mean that they save more overall. The counsel of experience therefore points—perhaps cautiously—*against* the hypothesis that workers irrationally undersave.

4) *Aggregated revealed preferences*: The wisdom of the crowds again suggests that actual savings behaviors are not irrational. The behaviors that BLE criticizes are not exceptional but rather typical. That is, the nudges and other manipulations that BLE analysts propose do not aim narrowly at outliers, who save much, much less than their peers, but rather at the middle of the savings distribution. Accordingly, BLE cannot point to normal savings rates as a way of framing lower rates as abnormal or irrationally too low. Put differently, even as the undersaving hypothesis has become conventional wisdom among BLE-influenced policymakers, it bears the burden of proof associated with countermanding the much broader conventional wisdom captured in normal

---

147. Beshears et al., *supra* note 35.

148. Thaler & Benartzi, *supra* note 122, at 173.

149. John Beshears, James J. Choi, David Laibson, Brigitte C. Madrian & William L. Skimmyhorn, *Borrowing to Save? The Impact of Automatic Enrollment on Debt*, 77 J. FIN. 403, 405 (2022). Moreover, yet another study finding that auto-enrollment increased savings concluded that the increase “appears to result from participant inertia and from employee perceptions of the default as investment advice,” and therefore not from the fact that opt-in savings rates irrationally undersatisfy normative preferences for saving. Brigitte Madrian & Dennis Shea, *The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior*, 116 Q.J. ECON. 1149, 1149 (2001). Other studies reach different results. See, e.g., Raj Chetty, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen & Tore Olsen, *Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark*, 129 Q.J. ECON. 1141 (2014) (finding, among Danish workers, that a 1 percent increase in mandatory contributions yielded a 0.8 percent increase in the total savings rate).

savings behavior. Thus, the wisdom of crowds also points *against* the hypothesis that workers irrationally undersave.

5) *Self-reported preferences*: People frequently report wanting to save more, as BLE analysts often point out. According to one study, “two-thirds of employees believe that they are saving too little and . . . one-third of these self-reported undersavers intend to raise their saving rate in the next two months.”<sup>150</sup> That said, “almost none of the employees who report that they intend to raise their saving rate in the next two months actually do so.”<sup>151</sup> Reports like these do nonetheless favor the undersaving hypothesis, but in the end only weakly. As the authors of the study admit, “[i]t’s not clear what subjects mean when they say that they save too little [and] [i]t’s also not clear what subjects mean when they say that they intend to raise their contribution rate in the next few months.”<sup>152</sup> The survey question at issue, which asked people to compare “actual” to “ideal” savings rates, only deepens the confusion.<sup>153</sup> Perhaps answers reflect the economist’s understanding of undersaving—that is, a savings rate that fails to maximize experienced well-being aggregated over time, given income and budget-constraints. But perhaps the answers merely reflect the folk-sense in which people worry about retirement and wish that they were richer. This interpretation receives some support from other prongs of the study—for example, the result that “getting someone to think about his or her own savings adequacy did not lead to any differential future behavior,”<sup>154</sup> which suggests that when people say that they undersave, they mean that they are not rich enough rather than that they are spendthrift. And in another study, Thaler and Benartzi admit as much, when they quote a pay consultant who explains that recommending high savings rates causes the “majority of workers” who “live paycheck to paycheck and can barely make ends meet” to do nothing to change their savings.<sup>155</sup> Workers’ self-assessments may *support* suggestions that they undersave, but this support is weak at best.

6) *Informed preferences*: We finally turn to informed preferences, as reflected by both more-informed decision-makers and the experts themselves. BCLM review the evidence that traditional financial education encourages more saving and reveal that it is quite weak.<sup>156</sup> But BLE experts do seem widely to share the view that workers do not save enough. This face of conventional expert opinion is captured, once again, by Thaler and Benartzi’s supposition that “[m]any countries are facing a retirement savings crisis”<sup>157</sup>—the proposition from which our investigation of savings adequacy set out. But any suggestion of

---

150. James J. Choi, David Laibson, Brigitte C. Madrian & Andrew Metrick, *Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance*, in 16 *TAX POLICY AND THE ECONOMY* 67, 70 (James Poterba ed., 2002).

151. *Id.*

152. *Id.* at 74.

153. *Id.* at 72.

154. *Id.* at 87.

155. Thaler & Benartzi, *supra* note 122, at 172.

156. Beshears et al., *supra* note 35, at 1793.

157. Benartzi & Thaler, *supra* note 38, at 1152.

deference to expert opinion in this context faces a challenge, namely the need to explain the grounds of expert belief. It cannot be sufficient that experts, applying a test that includes expert opinion, say that workers undersave, and that expert opinion is one of the factors that informs the question whether people save enough. This would be bootstrapping of the most insecure variety. Moreover, it is important to be clear about just which expertise those whose views constitute expert opinion must have. In order to inform the question what savings rate is optimal, the experts must have unusual knowledge of normative matters, concerning the conditions of human flourishing over the course of a life. This is very different from the types of empirical, statistical, and modelling expertise that BLE experts possess. Indeed, it is not clear that it is a subject that admits of expertise in the scholarly or academic sense at all. The final factor therefore also offers no substantial independent support to the undersaving hypothesis. Nevertheless, for the sake of argument, we count this factor as offering weak *support* for the undersaving hypothesis.

Overall then, two factors weakly favor the undersaving hypothesis, two weakly reject it, and two stand in equipoise—hardly a clear and convincing endorsement of the conclusion that workers badly undersave. And yet, behaviorally-influenced policymakers and applied BLE analysts commonly, indeed overwhelmingly, conclude that undersaving is real and dramatic. Often they justify these claims by cherry-picking just one or two of these six factors and citing incomplete and even tendentious accounts of the factors that they do cite. Thaler and Benartzi, for example, emphasize in one article that workers say that they undersave, while ignoring the ambiguity in what people might mean by this.<sup>158</sup> In another article—the one that begins by sounding the alarm about the “savings crisis”—they argue only that “the fraction of workers at risk of having inadequate funds to maintain their lifestyle through retirement is estimated to have increased from 31% to 53% from 1983 to 2010.”<sup>159</sup> But this measure sweeps up every worker who will experience any drop—no matter how small—in consumption at retirement.<sup>160</sup> And the years chosen distort the trend even more, by skewing savings upwards in 1983, when the S&P 500 was 25 percent higher than three years before and downwards in 2010, when it was 30 percent lower.<sup>161</sup> The collapse in housing prices following the Great Recession increases the distortion, as most households’ largest assets are their primary residences.<sup>162</sup> Such partial and tendentious evidence falls far short of the certain and dramatic

---

158. See Thaler & Benartzi, *supra* note 122, at 167.

159. Benartzi & Thaler, *supra* note 38, at 1152. This evidence is effectively an example of structural estimation, in which the model implies that consumption should be constant before and after retirement.

160. See Munnell, Webb & Golub-Sass, *supra* note 123.

161. Arna Olafsson & Michaela Pagel, *The Retirement-Consumption Puzzle: New Evidence from Personal Finances* 24-25 (Nat’l Bureau of Econ. Rsch. Working Paper No. 24405, 2018) (finding that, upon retirement, individuals reduce their spending in both work-related and leisure categories). The claims about the S&P 500 compare levels on June 1st.

162. See Munnell, Webb & Golub-Sass, *supra* note 123, at 5 (showing that half of the increase in “at-risk” households came from the recent housing price collapse).

conclusions that dominate the behavioral law and economics of retirement savings.<sup>163</sup>

BLE assessments of retirement savings, in practice, thus allow unstructured lists of considerations to yield partial arguments that focus opportunistically on factors spun to support favored conclusions. This pattern will be familiar to lawyers, who have long worried that multifactor tests encourage conclusory cherry-picking, which merely disguises intuitions behind a smoke-screen of false deliberation.<sup>164</sup> The risk here is that behavioral analyses of retirement savings lead an elite of “experts” to impose their (often eccentric) personal or group values, all under the guise of what is apt or even rational, in just the sort of turn to “intuition” that (as we earlier observed) Bernheim and Rangel reject. Economists—no matter how earnest and good-willed they are in their work—are not immune from this critique.

Certainly, the rich tend (for any number of reasons) to save at higher rates than the rest,<sup>165</sup> and bourgeois morality purports to rationalize this preference into a virtue.<sup>166</sup> As noted at this Article’s outset, economists, and especially the

---

163. Benartzi and Thaler also bemoan that roughly half of U.S. workers have no access to retirement plans through their jobs. But this is concerning only on the assumption that people undersave. If other savings (through Social Security, for example, or home equity) are adequate, then the lack of employer-administered pensions becomes evidence that such savings are not needed and therefore not much wanted.

164. See, e.g., T. Alexander Aleinikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L.J. 943 (1987) (explaining different forms of balancing tests and providing both external and internal critiques of balancing); James Wilson, *The Morality of Formalism*, 33 UCLA L. REV. 431 (1985) (critiquing multifactor balancing tests for being unpredictable and malleable using examples from the death penalty, abortion, and equal protection contexts).

165. Generally, the positive correlation between income level and saving is well established. See, e.g., Patti Fisher & Sophia Anong, *Relationship of Saving Motives to Saving Habits*, 23 J. FIN. COUNSELING & PLAN. 63, 69 (2012) (noting that lower-income households do not save or save significantly less than higher-income households). Furthermore, most studies tend to indicate a positive relationship between education and views on savings. For example, one study indicated that years of education had a positive effect on attitudes toward savings (as measured by both opportunity to participate in a retirement preparation program and actual participation in such a program). Scott Beck, *Retirement Preparation Programs: Differentials in Opportunity and Use*, 39 J. GERONTOLOGY 596, 600 (1984). Another study also found that education was associated with savings: compared to those who did not graduate from high school, those who did were 33% more likely to save regularly, and those with a graduate degree increased the odds of saving regularly by 34%. The savings behavior of those with some college was not significantly different from those who did not graduate high school. Celia Hayhoe, Soo Hyun Cho, Sharon A. DeVaney, Sheri Lokken Worthy, Jinhee Kim & Elizabeth Gorham, *How Do Distrust and Anxiety Affect Saving Behavior?*, 41 FAM. & CONSUMER SCI. RSCH. J. 69, 77-78 (2012). The study concludes that understanding people’s differing financial attitudes can be helpful in developing intervention programs to improve financial management and savings behaviors. *Id.* at 81.

166. See, e.g., Fisman et al., *supra* note 7, at 1300-04 (finding that elite subjects were “substantially more efficiency-focused” than were non-elites); see also Adam Grant, *Does Studying Economics Breed Greed?*, PSYCHOL. TODAY (Oct. 22, 2013), <https://www.psychologytoday.com/us/blog/give-and-take/201310/does-studying-economics-breed-greed> [<https://perma.cc/QPQ5-UR3R>] (citing academic work finding differing preferences between economists and non-economists and that studying economics can change students’ preferences). We note that these moralisms can valorize the poor as well as besmirch them. The UK’s Behavioural Insights Team, for example, observed that “low-income individuals may be less susceptible to certain framing effects,” compared to rich people who “were willing to pay more for a beer in the context of an expensive hotel compared to in a grocery store [compared to] low-income groups [who] were more consistent across contexts in what they would be willing to pay for a beer.” Gandy et al., *supra* note 129, at 30-31 (citing Richard Thaler, *Mental Accounting and Consumer Choice*, 4 MKTG. SCI. 199 (1985)). The Team concluded that that “[t]his appears to be because [poor

elite scholars who dominate BLE analysis of retirement savings, are undoubtedly rich.<sup>167</sup>

More generally, as noted earlier, economists have distinctive demographics, embrace distinctive values, and hold distinctive preferences, even compared to other elites. The under-representation of women and racial minorities in economics is striking.<sup>168</sup> And survey evidence suggests that economists and non-economists have different policy preferences, even after the non-economists have been informed of relevant facts, whether because of selection into economics or treatment through economics training.<sup>169</sup> Economists even insulate themselves from other intellectuals, with far higher rates of within-field citation than other social sciences<sup>170</sup> and greater distaste for interdisciplinary knowledge.<sup>171</sup> And there is evidence that economists import these values into their scholarly work, even when they aim to be purely descriptive, and possibly without being aware of what they are doing.<sup>172</sup> Much of this reflects the salutary fact that economists hold their professional values sincerely and apply them consistently across their entire lives. But it hardly makes their intuitions a reliable guide to the preferences and welfare of others. Instead, all these considerations suggest that BLE-driven savings policies might reflect the extraordinary preferences of behavioral economists rather than the normative preferences and well-being of the citizens whom the policies would govern.

Nor is this a purely theoretical concern. Economists have not been shy about moralizing based on such parochial ethics, as when Greg Mankiw (sometime chair of Harvard's economics department) criticized Supreme Court nominee

---

people's] financial situation focuses their attention on spending trade-offs (i.e., what else could that money buy me?) and this gives them a stable (and more 'rational') framework for judging value-for-money." *Id.* (citing Anuj K. Shah, Eldar Shafir & Sendhil Mullainathan, *Scarcity Frames Value*, PSYCHOL. SCI. (2015)). Of course, another possibility is simply that the rich and poor place different values on drinking in the environments provided by posh hotels.

167. See *supra* note 8 and accompanying text.

168. NAT'L SCI. FOUND., *supra* note 12.

169. Paola Sapienza & Luigi Zingales, *Economic Experts Versus Average Americans*, 103 AM. ECON. REV. 636, 637-39 (2013); see Christopher Johnston & Andrew Ballard, *Economists and Public Opinion: Expert Consensus and Economic Policy Judgments*, 78 J. POL. 443, 449-51 (2016).

170. Marion Fourcade, Etienne Ollion & Yann Algan, *The Superiority of Economists*, 29 J. ECON. PERSP. 89, 93 (2015) (showing 81% within-field citation rates as compared to between 52% and 59% for sociology, anthropology, and political science (citing JERRY JACOBS, IN DEFENSE OF DISCIPLINES: INTERDISCIPLINARITY AND SPECIALIZATION IN THE RESEARCH UNIVERSITY 82 (2015))).

171. *Id.* at 95 (showing only 42% agreement among economists with the proposition that "In general, interdisciplinary knowledge is better than knowledge obtained by a single discipline," as compared to between 60% and 87% in other fields including sociology, political science, psychology, finance, and history).

172. See Zubin Jelveh, Bruce Kogut & Suresh Naidu, *Political Language in Economics 1* (Columbia Bus. Sch. Rsch. Paper No. 14-572018), <https://ssrn.com/abstract=2535453> [<https://perma.cc/MZN4-CV5L>] (finding suggestive evidence that political ideology influences results of economic research—namely, using natural language processing and machine learning to show a correlation between the observed political behavior and language of economists on the one hand and their empirical results on the other); see also Adam Bonica, Adam Chilton, Kyle Rozema & Maya Sen, *The Legal Academy's Ideological Uniformity*, 47 J. LEGAL STUD. 1, 32-34 (2018) (showing that law professors have political views that are non-representative of lawyers overall).

Sonia Sotomayor as spendthrift, only to backtrack when it was pointed out that her career as a judge entitled her to a substantial federal defined benefit pension.<sup>173</sup> And the myriad BLE-inflected warnings of a retirement savings crisis generalize this personal moralizing into public policy advice. Ordinary people, by contrast, seem not to share this concern for undersaving, not just in their private conduct as workers but also in their public conduct as citizens. The retirement savings crisis that dominates applied BLE scholarship on savings and also the work of technocratic behavioral science teams<sup>174</sup> casts no substantial shadow over electoral politics. Increasing retirement savings has strikingly little salience among either Democratic or Republican voters.<sup>175</sup>

One of us (Markovits) also worries that behavioral interventions might not just impose elite values but also enable elites to launder their own self-interest through nudges that they frame as other-regarding. Retirement savings illustrate this risk. Thaler and Benartzi, in promoting their Save More Tomorrow™ plan for increasing retirement savings, observe that a company adopting the plan was motivated in part by the fact that federal non-discrimination rules (which limit how much pension plans may favor elite over rank-and-file workers) caused low savings rates among middle-class workers to limit how much executives could contribute to their retirement plans.<sup>176</sup> And in another study of opt-out savings plans, Choi, Laibson, Madrian, and Metrick similarly admit that “[t]he interest of many companies in automatic enrollment has stemmed from their persistent failure to pass . . . nondiscrimination rules that apply to pension plan provision[s]” so that the firms had to “make either *ex post* . . . contribution refunds to highly compensated employees or retroactive company contributions on behalf of non-highly compensated employees.”<sup>177</sup> Moreover, interventions that increase retirement savings also increase the fees captured by the investment and pension firms that inevitably manage and administer the savings. Finally, even state governments benefit from adopting “auto-IRAs,” which automatically default workers into retirement savings plans, because workers with lower savings “tend to start taking Social Security benefits the moment they become eligible, which means they get the smallest benefit they could and may also come to rely on other government services for low-income individuals.”<sup>178</sup> Motives, of

---

173. Greg Mankiw, *SCOTUS Nominee Is a Spender*, GREG MANKIW'S BLOG (May 26, 2009), <http://gregmankiw.blogspot.com/2009/05/scotus-appointee-is-spender.html> [https://perma.cc/Q4TQ-6V6M].

174. OFF. OF SCI. & TECH. POL'Y (2015), *supra* note 134; OFF. OF SCI. & TECH. POL'Y (2016), *supra* note 128.

175. Though members and leadership of both parties seem concerned with the protecting existing social welfare programs like Social Security and Medicare, there appears to be little concern with expanding retirement savings yet further. See Joy Wilke & Frank Newport, *Democrats and Republicans Differ on Top Priorities for Gov't*, GALLUP (Jan. 28, 2014), <https://news.gallup.com/poll/167084/democrats-republicans-differ-top-priorities-gov.aspx> [https://perma.cc/4SNV-BLL8].

176. See Thaler & Benartzi, *supra* note 122, at 171.

177. Choi et al., *supra* note 150, at 75.

178. Bailey McCann, *States Take Aim at People with No Retirement Plan*, WALL ST. J. (July 7, 2020, 9:38 AM EST), <https://www.wsj.com/articles/states-take-aim-at-people-with-no-retirement-plan-11593945474> [https://perma.cc/4BU2-KLAX].

course, are hard to discern. But it is difficult, Markovits notes, to exclude the possibility that elite conviction that a savings crisis exists is a case of the moralist's laundering self-interest as high-minded other-regard. As Bernard Williams once observed, "[h]e would be better off dead' can be said for many dubious reasons: the most dubious of which is that we would be better off if he were dead."<sup>179</sup> The same goes for savings.

In any case, however well-intentioned, well-trained, and intelligent BLE scholars may be, the concerns of bias remain. The reason is that, in the world that BLE's descriptive triumphs reveal, the problem of determining optimal savings cannot be solved using the ideas and methods that economics develops. Expert opinion here simply outstrips the technical capabilities that constitute economic expertise.

### C. *The Cost of Manipulations that Increase Savings*

These errors are not merely abstract or academic, and they are far from costless. The view that workers do not save enough and that nudges should be used to remedy the situation has been deployed by governments through state "auto-IRAs," encouraged by Obama Administration regulations,<sup>180</sup> which automatically default workers into retirement savings plans. Since BLE-inflected savings policy came into vogue (specifically, following a 2006 study by Brookings and Heritage economists<sup>181</sup>), at least 10 states have implemented policies that manipulate workers in this way.<sup>182</sup> These policies seem to have been adopted without a single publicly available analysis of their impact on low- or middle-income families, despite the fact that many such families need immediate access to the income that they earn in order to pay for basic necessities.

BLE's boosters emphasize that nudges are not commands and insist, with Richard Thaler, that "[t]here is no coherent argument against . . . state plans"<sup>183</sup> that use behavioral effects to increase savings. Indeed, a leading concern among BLE scholars working on retirement savings is that nudges that default people into retirement plans might, if default contribution rates are low, have the unwanted effect of reducing savings over the long run: the stickiness of the low rates could block decisions to save at higher rates that workers, left to their own

179. BERNARD WILLIAMS, *ETHICS AND THE LIMITS OF PHILOSOPHY* 48 (1986).

180. Yuka Hayashi, *U.S. Nudges States to Help Private-Sector Workers Save for Retirement*, WALL ST. J. (Aug. 25, 2016, 6:46 PM EST), <https://www.wsj.com/articles/u-s-rule-encourages-states-to-offer-retirement-plans-for-private-sector-workers-1472149529> [<https://perma.cc/ZH76-GSGV>].

181. J. Mark Iwry & David John, *Pursuing Universal Retirement Security Through Automatic IRAs*, RET. SEC. PROJECT (July 1, 2009), [https://www.brookings.edu/wp-content/uploads/2016/06/07\\_automatic\\_ira\\_iwry.pdf](https://www.brookings.edu/wp-content/uploads/2016/06/07_automatic_ira_iwry.pdf) [<https://perma.cc/XSY8-NZVC>].

182. McCann, *supra* note 178.

183. Richard Thaler (@R\_Thaler), TWITTER (Mar. 28, 2017, 4:53 PM), [https://twitter.com/R\\_Thaler/status/846827550093455360?s=20](https://twitter.com/R_Thaler/status/846827550093455360?s=20) [<https://perma.cc/366Q-9X4X>]; see Richard Thaler, *State I.R.A. Plans Are Ready, If Congress Doesn't Interfere*, N.Y. TIMES (Mar. 3, 2017), <https://www.nytimes.com/2017/03/03/upshot/state-ira-plans-are-ready-if-congress-doesnt-interfere.html> [<https://perma.cc/5Q9K-HJJN>].

devices, would otherwise make. Madrian and Shea thus worry that “[a]utomatic enrollment appears to be a win-lose approach to changing 401(k) savings behavior.”<sup>184</sup> The “win aspect” is the increase in savings from the increase in participation, while the “lose aspect” is the decrease in savings from participant inertia in the contribution rate.<sup>185</sup> And they respond to this concern by doubling down on behavioral manipulations, proposing that “[t]o turn automatic enrollment from a win-lose proposition to a win-win proposition, employers must find ways to move employees into higher contribution rates and more aggressive investment strategies.”<sup>186</sup>

But these arguments move much too quickly, on many fronts, and the expansion of state auto-IRA plans illustrates the costs of manipulations to increase savings.

First, the distinction between nudges and commands cannot sustain the moral confidence that remarks such as Thaler’s project. A vigorous empirical debate asks whether behavioral manipulations produce the durable impacts on savings that boosters claim<sup>187</sup> or have only short-term effects which dissipate as workers opt out of default plans or adjust other borrowing in light of staying in them.<sup>188</sup> But insofar as BLE savings policies do induce a material and durable change in savings behavior (using means that bypass workers’ conscious choices), then the distinction between nudges and commands becomes, as others have noted, a distinction without a difference.<sup>189</sup> Certainly it enjoys no talismanic power to absolve BLE analysts from the duty to provide a moral justification for their interventions.

Second, and crucially, BLE policies cannot provide the required moral justifications. Manipulations that successfully encourage increased savings have costs. They risk inducing workers to save more than their true preferences recommend, pushing savings towards the levels set according to elite BLE analysts’ distinctive preferences rather than according to the worker-savers’ own, more typical ones. This would reduce workers’ lifetime-experienced well-being. Moreover, there is good reason to suspect that the harm to well-being will fall disproportionately on those who are already worst off. In our unequal society,

184. Madrian & Shea, *supra* note 149, at 1185.

185. *Id.*

186. *Id.*

187. See, e.g., Thaler & Benartzi, *supra* note 122, at 173, 179 (explaining that automatic enrollment plans “can produce dramatic increases in saving rates”); Benartzi & Thaler, *supra* note 38, at 1152 (similar); James Choi, David Laibson, Brigitte C. Madrian & Andrew Metrick, *For Better or for Worse: Default Effects and 401(k) Savings Behavior*, in PERSPECTIVES ON THE ECONOMICS OF AGING 81, 81-83 (David A. Wise, ed., 2001) (similar); Madrian & Shea, *supra* note 149, at 1149 (similar).

188. See, e.g., Choukhmane, *supra* note 126, at 1, 14 (showing based on data from 401(k) plan that “the long-term effect of auto-enrollment on wealth is negligible except” for low-income workers and that, though “[a]utoenrolled workers initially contribute more to the retirement savings plan,” for most of the income distribution, “the nonautoenrolled employees catch up over three years and contribute a similar total amount to the 401(k) plan as the autoenrolled workers”); Beshears et al., *supra* note 149. One other possibility is that where auto-enrollment sets low default savings rates, the stickiness of the nudge might over time itself suppress savings. See Madrian & Shea, *supra* note 149, at 1185 (providing such evidence).

189. Bubb & Pildes, *supra* note 34; see also Gregory Mitchell, *Libertarian Paternalism is an Oxymoron*, 99 NW. L. REV. 1245 (2004).

the poor and even the lower middle class have very little room for saving. The Consumer Expenditure Survey reports that households in the bottom fifth of the income distribution devote 172 percent of after-tax income to housing, food (excluding alcohol), transportation, and healthcare.<sup>190</sup> The next poorest fifth of households spend 100 percent of their income on housing, food, transportation, and healthcare.<sup>191</sup> For fully 40 percent of households, therefore, manipulations that increase savings come at the cost of decreasing current consumption of obvious necessities, which is especially concerning because many working people have elevated needs for necessities as they are raising children at this stage of their lives, unlike when they are receiving Social Security after retirement.<sup>192</sup> At the same time, poorer households get a much smaller tax benefit for each dollar of retirement savings than richer ones because they face lower marginal tax rates.<sup>193</sup>

#### D. Other Examples

The retirement savings example is not exceptional. BLE-inflected policy displays analogous pathologies in other arenas—in which elite policymakers, making insufficiently grounded assumptions about rationality, risk imposing their own preferences under the guise of enlightened paternalism.

Consider, for example, policy on geographic mobility. A significant behavioral literature is developing on efforts to encourage people to move from so-called low-“opportunity” places to higher-“opportunity” places.<sup>194</sup> Raj

190. *Table 1101. Quintiles of Income Before Taxes: Annual Expenditure Means, Shares, Standard Errors, and Coefficients of Variation, Consumer Expenditure Survey, 3rd Quarter 2017 Through 2nd Quarter 2018*, BUREAU LAB. STAT. (May 19, 2019), <https://www.bls.gov/cex/tables/mid-year/mean-item-share-average-standard-error/cu-income-quintiles-before-taxes-2017-2018.pdf> [https://perma.cc/BND8-X6CK]; see also Taylor Cranor, *Opting for Active Choice: Addressing the Distributive Consequences of Auto-IRAs* (unpublished student paper) (on file with authors); Andrew Biggs, *Stop Pushing Poor People to Save More for Retirement*, AM. ENTER. INST. (Sept. 12, 2019), <https://www.aei.org/articles/stop-pushing-poor-people-to-save-more-for-retirement> [https://perma.cc/3J5N-KHHT].

191. *Table 1101, supra* note 190.

192. It is unsurprising, but worth making explicit, that the optimal wealth accumulation for poor households, conditional on their remaining poor, is very small. Low savings are a rational response and thus a symptom of poverty. See, e.g., Scholz, Seshadri & Khitatrakun, *supra* note 126, at 624 (“The optimal wealth target for the median households in the lowest decile of the lifetime earnings distribution is very low, at \$2,050 (including housing wealth).”).

193. In 2021, a married couple earning less than \$25,100 faces a federal marginal income tax rate of 0%, and a married couple earning less than \$45,000 faces a rate of only 10%. Treas. Reg. § 601.602 (1981); Rev. Proc. 2020-45, 2020-46 I.R.B. 1016, 1022 (providing for a standard deduction of \$25,100 and a 10% tax bracket for the next \$19,900). For a similar point, see Cranor, *supra* note 190.

194. Gharad Bryan, Shyamal Chowdhury & A. Mushfiq Mobarak, *Seasonal Migration and Risk Aversion*, DEP’T OF LABOR (Dec. 16, 2011), [https://ecommons.cornell.edu/bitstream/handle/1813/77973/ILAB\\_Seasonal\\_Migration\\_and\\_Risk\\_Aversion.pdf](https://ecommons.cornell.edu/bitstream/handle/1813/77973/ILAB_Seasonal_Migration_and_Risk_Aversion.pdf) [https://perma.cc/A747-49K8] (finding positive results from loaning money to farmers in Asia and Africa, thus concluding that “encouraging seasonal migration towards where jobs are appears to be a sensible complementary policy to experiment with”); Peter Bergman, Raj Chetty, Stefanie Deluca, Nathaniel Hendren, Lawrence F. Katz & Christopher Palmer, *Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice*, [https://opportunityinsights.org/wp-content/uploads/2019/08/cmto\\_paper.pdf](https://opportunityinsights.org/wp-content/uploads/2019/08/cmto_paper.pdf) [https://perma.cc/3C69-NEZS]; Crystal C. Hall, Martha M. Galvez & Isaac M. Sederbaum, *Assumptions About Behavior and*

Chetty, for example, proposes that people might irrationally stay in place on account of present bias, inadequate information, projection bias, and a scarcity mentality that can “amplify individuals’ focus on immediate needs.”<sup>195</sup> Policies based on such thinking are beginning to be adopted. Chetty and his coauthors have now partnered with cities to randomize interventions that encourage people to “move to opportunity.”<sup>196</sup>

We take no view on the ultimate question of whether people irrationally forswear economic opportunities that they might grasp, if only they could overcome behavioral obstacles to moving. We observe only that economics experts may be particularly likely to view this policy question differently from non-professionals. Economists disproportionately value material wealth, and elites (especially academic elites) disproportionately value professional prestige and are willing to move long distances in pursuit of it. Others may instead value community connections, family, and rootedness in place.<sup>197</sup> The risk of allowing policy elites to take a “pragmatic perspective” to policymaking under the guise of doing neutral science is that this may lead them to impose their personal values on everyone else.

A second example concerns education policy. A rising literature aims to display the enormous, even tragic, costs incurred when high-achieving, low-income students do not attend top universities.<sup>198</sup> Once again, work documenting a descriptive pattern has led to normative work proposing behavioral interventions designed to nudge choices away from revealed and towards normative preferences, and in this way to increase experienced well-being.<sup>199</sup>

---

*Choice in Response to Public Assistance: A Behavioral Decision Analysis*, 1 POL’Y INSIGHTS FROM BEHAV. & BRAIN SCI. 137 (2014); Andrea Lannom, *What’s Influencing People to Move Out of West Virginia*, TIMES W. VA. (Dec. 27, 2017), [https://www.timeswv.com/news/what-s-influencing-people-to-move-out-of-west-virginia/article\\_875b3b6a-eac6-11e7-93f0-73fe86474029.html](https://www.timeswv.com/news/what-s-influencing-people-to-move-out-of-west-virginia/article_875b3b6a-eac6-11e7-93f0-73fe86474029.html) [https://perma.cc/N37S-Q7P4] (noting that West Virginia ranked low on the AARP’s Livability Index on categories of health and access to life, work, and play, resulting in the movement of educated people to higher-opportunity areas).

195. Chetty, *supra* note 27, at 21-22.

196. See Chetty et al., *supra* note 194, at 1.

197. See, e.g., Larissa MacFarquhar, *In the Heart of Trump County*, NEW YORKER (Oct. 3, 2016), <https://www.newyorker.com/magazine/2016/10/10/in-the-heart-of-trump-county> [https://perma.cc/XU8F-43MZ] (detailing many people’s attachments to their home town and home state).

198. Caroline M. Hoxby & Christopher Avery, *The Missing “One-Offs”: The Hidden Supply of High-Achieving, Low Income Students* 1 (Nat’l Bureau of Econ. Rsch., Working Paper No. 18586, 2012) (finding that the majority of high-achieving, low-income students do not apply to any selective colleges despite being well-qualified for admission); see also Shereen Marisol Meraji, *Why Many Smart, Low-Income Students Don’t Apply to Elite Schools*, NPR (Mar. 16, 2015, 4:16 PM EST), <https://www.npr.org/2015/03/16/393339590/why-many-smart-low-income-students-dont-apply-to-elite-schools> [https://perma.cc/U5FK-AK6F] (explaining the possible reasons high-achieving, low-income students do not apply to selective universities, including distance and lack of knowledge).

199. E.g., Keith O’Brien, *Nudging: Yes, Higher Ed, It Does Help Student Outcomes*, SIGNALVINE (Aug. 22, 2018), <https://www.signalvine.com/nudge-technology/nudging-student-outcomes> [https://perma.cc/4NRT-CQ73] (arguing that higher-education professionals should use “nudging” to persuade students to matriculate). A second set of policies encourages colleges to recruit less advantaged students more widely and effectively. See Rodney J. Andrews, Scott A. Imberman & Michael F. Lovenheim, *Recruiting and Supporting Low-Income, High-Achieving Students at Flagship Universities*, 74 ECON. EDUC. REV. 1, 16-18 (2020); Hoxby & Avery, *supra* note 198. These policies do not raise the

As before, we take no view here concerning whether it is good or just to increase the rates at which high-achieving, low-income students enroll at elite colleges. (Although we note that one of us has elsewhere argued for using state coercion dramatically to increase such enrollments.<sup>200</sup>) We observe, only, that education policy raises profound moral questions, which cannot be adequately addressed by consulting intuition, and that here as elsewhere, elite economists likely hold idiosyncratic preferences and values. The financial costs and benefits of attending college (and the distribution of these costs across different types of students and colleges) may be straightforwardly and uncontroversially measured. But other costs and benefits quickly engage reasonable moral disagreements. As with geographic mobility, many students, especially from working-class backgrounds, may value staying near their homes, families, and communities—both geographically and culturally. Some students, while academically gifted, may not like school and may recoil at the prospect of many years of schooling, running deep into adulthood. Gainsaying these preferences may have substantial benefits, but it also carries real costs: an academic from a working-class background once observed, on returning home to his original community, that “I feel like I have changed sides in some very important game.”<sup>201</sup> More systemically, promoting university education comes (in a world of resource constraints) at the cost of neglecting vocational and on-the-job training. Elite BLE scholars, most of whom come from professional backgrounds and virtually all of whom have a taste for education and scholarship, generally do not share preferences that are skeptical of schooling; and some may find them difficult to credit. But they should not be dismissed as irrational—as reflecting “present bias,” a failure to consider future selves, impoverished cognition, or some other bias.

These examples may be further multiplied. Indeed, the scope for “easy cases” that avoid hard normative questions seems quite limited. We do not deny that some interventions—such as removing a duplicative question on a form for a benefit that all agree an applicant should receive—may not be vulnerable to our criticisms. But other seemingly benign interventions turn out, on reflection, quickly to involve contestable moral judgments. For example, even information interventions—say, about the benefits of vaccines—present tradeoffs. Information interventions take resources, including money spent getting people’s attention and the attention itself, which inevitably distracts people from other matters. To know how much to spend and how much attention to capture requires knowing the benefits of the vaccines relative to the costs to individuals of receiving them, which immediately gets policymakers back into the normative thicket.

---

concerns that we emphasize, as they seek to expand student options rather than to manipulate student choices.

200. See DANIEL MARKOVITS, *THE MERITOCRACY TRAP* 276-79 (2019).

201. John Sumser, *Working it Out: Values, Perspectives, and Autobiography*, in *THIS FINE PLACE SO FAR FROM HOME: VOICES OF ACADEMICS FROM THE WORKING CLASS* 304 (C.L. Barney Dews & Carolyn Leste Law eds., 1995).

Once again, none of our discussions—not even the more elaborate account we have given of retirement savings—can sustain conclusions about what policy should be. Rather, we claim only that—in these cases as in myriad others—BLE-inspired policy work, rather than following technocratic logic about what rational people would necessarily choose, instead risks reflecting the contingent and contestable moral prejudices of elite BLE scholars. Narrow and even eccentric cultural biases risk displacing careful and systematic moral argument. This happens not because BLE scholars are unscrupulous or lazy, but rather, as the close analysis of retirement savings reveals, because BLE's descriptive revolution makes normative work immensely difficult.

### III. Towards a Democratic Behavioral Law and Economics

Public policy typically raises questions that concern both expert and everyday matters. The optimal level of retirement savings, to return to the familiar example, depends on both technical questions about how to administer and grow a nest-egg and commonsense questions about the relative values people place on present gratification and comfort in old age. A long and distinguished tradition in legal thought aspires to integrate technocratic expertise and ordinary beliefs, desires, and choices and, in this way, to produce policy outcomes that are better and more legitimate than either expert or everyday inputs can achieve alone.

BLE's descriptive accomplishments massively complicate this integrative project by casting doubt over ordinary people's capacities to accurately and effectively promote their own interests—a capacity they have long been simply assumed to possess. At the same time, BLE's disciplinary roots in traditional law and economics lead astray normative BLE's efforts to manage the complications. In particular, these roots invite the normative mistakes that we have documented, in which BLE scholars risk imposing their own personal preferences and biases on the general population.

We revisit broader traditions of integrating expert and ordinary opinion in order to clarify the genealogy of BLE's normative errors and to chart a path out of the moral morass in which BLE currently finds itself. We frame our approach—as we have done throughout the paper—principally in terms of finding the policy that maximizes individuals' well-being (although we note that a parallel framing in terms of freedom and dignity is also available). The new path emphasizes a discursive engagement between expert and everyday opinion, in which BLE scholars and technocrats assist ordinary people in their deliberations, rather than directing or supplanting ordinary choices. For this reason, we call the approach that we introduce *democratic behavioral law and economics*.

This democratic approach stands against the *technocratic* approach to BLE that dominates the field today, including in the methods and examples that we have already criticized. We begin by describing the technocratic and democratic

approaches to BLE in turn, drawing out a series of sharp contrasts between them, which we summarize in Table 1. Three contrasts—which emphasize distinct but complementary perspectives on the problem of integrating expert and ordinary decision-making—stand out especially. First, the technocratic approach imposes a firm division of labor between expert opinion and ordinary choices, while the democratic approach brings the two into a shared, discursive engagement. Second, the technocratic approach debiases ordinary choosers by purifying their preferences to remove the irrationalities that descriptive BLE identifies, while the democratic approach debiases by helping ordinary choosers achieve self-conscious mastery over their behavioral tendencies. And third, the technocratic approach puts technical experts in control of public policy, while the democratic approach deploys experts as assistants, or tech-support, to empower ordinary people. Together, these contrasts illuminate the democratic approach and also highlight the reasons that can make it better than technocratic BLE.

Finally, we end this Part with an example of democratic BLE applied to retirement savings.

Although democratic methods in law and economics could well extend beyond areas with behavioral biases, we limit our discussion here to BLE, as this is the subject that both necessitates and allows the move beyond the purely technocratic approach.

Table 1: Contrasting Technocratic and Democratic Approaches to BLE

	Technocratic	Democratic
Model for integrating expert and ordinary opinion	Division of labor	Discursive engagement
Approach to debiasing	Purifying: exclude biases and identify how unbiased people would choose	Deliberative: teach people about their potential biases and invite them to choose afresh in that knowledge
Relationship between experts and ordinary people	Expert as paternalistic manager	Expert as tech-support

*A. Technocratic Behavioral Law and Economics*

BLE has long taken a “technocratic” approach to integrating its expertise into policymaking. This technocratic approach begins with a strict division of labor in making policy recommendations: ordinary people form opinions about what will maximize their well-being and, based on these opinions, make choices

in their everyday lives; experts, for their parts, recommend policies that frame and, where necessary, manipulate (or “nudge”) ordinary choices, in order to ensure that these choices reliably track normative preferences, or true well-being. We place this technocratic approach into a broader intellectual context, explain the problems that it involves, and connect these problems to our earlier criticisms of the current state of normative BLE.

BLE’s current troubles are connected to its roots in traditional law and economics. There, experts design the legal rules—for example, the doctrines that make up the law of property, tort, and contract—that together establish the market frameworks within which ordinary people invest, trade, and consume. The experts deliberate among themselves about what rules are best: scholarly journals are filled with arguments about which doctrines—evaluated either individually or together in ways that produce a certain type of market—will promote efficiency and justice. Ordinary people, by contrast, contribute choices rather than arguments—they participate through will, not reason, simply choosing how to invest, trade, and consume, without any requirement or even expectation that they justify or explain their decisions.

This stark division of labor between experts and ordinary people, moreover, reflects normative commitments rather than just habits or facts. As we observed earlier, the ideal of rational choice, embraced by traditional law and economics, turns the division of labor into a principle: experts may fix the frames within which people choose and may argue among themselves about how to construct these frames; but it is not for experts, or indeed for anyone, to question ordinary people’s choices within the frames. By the same token, it is not for ordinary people—at least not based on their brute preferences—to challenge expert judgment about what efficient property, tort, and contract law looks like. At some places and times—notably, in the United States during the *Lochner* era—this deference to expertise has been constitutionalized, to insulate the expert-driven private law that constructs markets from even concerted democratic politics. But even where democratic forces can influence the rules that shape the market, experts retain a substantial measure of independence, captured in their insistence on the distinction between law and choices made under law. This division of labor—and the insulation of technical expertise from meddling by ordinary people—makes traditional law and economics technocratic.

BLE’s descriptive findings confront this approach to integrating expert and ordinary opinion with a potent new challenge. Conventional accounts of expertise understand technical training as adding something—knowledge or skill—to ordinary reason. At the same time, these accounts treat ordinary reason as adequate within its domain—for example, when ordinary people make their everyday decisions. BLE, by contrast, elaborates a deep and pervasive skepticism of ordinary reason, which puts the assumption about the adequacy of ordinary reason under immense strain. Behavioral effects undermine the reliability of ordinary beliefs and choices even in everyday situations, where no technical questions are at issue. Where revealed preferences diverge from normative preferences, ordinary people, even in everyday contexts such as the

choice between spending and saving, fail effectively to promote their own welfare. BLE's descriptive insights, therefore, reveal that ordinary choices are often (pervasively?) untrustworthy, introducing a new kind of obstacle to the project of integrating expert and ordinary opinion.

Descriptive BLE's skepticism of everyday competence leaves normative BLE with little reason to give ordinary choices a prominent role in policy. Insofar as revealed preferences depart from normative preferences, assumptions about rationality are undermined; and the reasons that led traditional law and economics to embrace ordinary market choices (made within the technocratically constructed market frame) dissolve. At the same time, traditional law and economics's division-of-labor approach provides normative BLE with no model for a discursive or deliberative engagement between expert and ordinary opinion. Because experts and ordinary people act in separate, distinct, and fundamentally different realms, their competences do not overlap or even address each other.

When descriptive BLE undermines the competence of ordinary choosers even within their traditional realm, technocrats will naturally expand their sphere of influence.<sup>202</sup> Taken together, the intellectual forces just outlined encourage normative BLE not only to embrace technocracy but also to reject democracy, just as it has done. Technocratic experts, that is, decide what unbiased choices are and then, using BLE's descriptive insights, design choice architectures to induce irrational (and often unwitting) ordinary people to make the "correct" choices. Where descriptive BLE undermines rational choice theory, normative BLE steps in paternalistically to fill the newly opened void.

The techniques that normative BLE uses to identify normative preferences—to know what experienced well-being consists in—provide technocratic BLE with the methods it needs to implement its paternalistic agenda. Each of these techniques seeks, adopting a term from the behavioral lexicon, to debias<sup>203</sup> ordinary, irrational people, specifically by purifying their revealed preferences to excise the various irrationalities that descriptive BLE identifies. Structural estimation aspires to model and quantify the effects of a behavioral bias in order then to remove those effects and recover unbiased normative preferences from actual behavior. In this way, it enables the BLE analyst to reconstruct the choices that rational agents would have made if they had behaved as if they were free from their biases. Modified revealed preferences analysis pursues an analogous approach, only now replacing structural estimation's theoretical effort to isolate biases in all choices and agents with an empirical effort to segment the population into biased and unbiased choices and agents. The BLE analyst then once again uses this segmentation to quantify the effects of biases and to reconstruct rational choices, made as if the biases had been excised. Finally, hedonics seeks to excise biases by circumventing

---

202. See, e.g., Kristen Underhill, *Broken Experimentation, Sham Evidence-Based Policy*, 38 YALE L. & POL'Y REV. 150 (2020).

203. See, e.g., Christine Jolls & Cass Sunstein, *Debiasing Through Law*, 35 J. LEGAL STUD. 199, 199 (2006) (describing "attempt[s] to debias through law by steering people in more rational directions").

preferences and choices entirely and relying instead on direct reports of well-being, so that biases are removed by being, in effect, ignored. For all their differences, therefore, all three approaches share a commitment to defeating biases by isolating and removing their effects—to purifying preferences, as we say.

The medley approach that BLE-inflected policy analysis deploys in practice again imports this general division-of-labor attitude from the theoretical traditions that it draws on and embraces the paternalistic idea that debiasing requires purifying preferences of the irrationalities that descriptive BLE reveals. When government behavioral insights teams construct manipulations designed to nudge people towards choices that reflect their normative preferences, the technocrats who constitute these teams are assumed to deploy debiased expert judgment on behalf of those whom they manipulate, against their revealed preferences, into unwittingly rational choices. The guiding idea behind these interventions is that experts—on account of the training that technocrats have received and the methods (including the pure methods just described) that technocrats deploy—can identify, isolate, and excise the biases that pollute ordinary choices, to serve the normative preferences and maximize the experienced well-being of ordinary people. Technocratic BLE reduces the cost and increases the effectiveness of debiasing. Only expert technocrats need to learn about the biases; and they can then spread the benefits of debiasing across the whole population.

Once again, BLE's intellectual roots in traditional law and economics—and in this tradition's division-of-labor approach to integrating expert and ordinary opinion—make it natural for normative BLE to pursue a paternalistic approach to debiasing and to embrace the technocratic management that paternalism demands. Descriptive BLE undermines rational choice theory, and when one side of a division of labor is stripped of its authority, the other naturally expands into the now-open field.

Natural, but not justifiable. As we have shown, normative BLE remains ill-equipped to occupy the field into which it has advanced. Expert-driven debiasing can purify revealed preferences of irrationalities only insofar as the experts can identify and isolate the irrationalities at issue and know what normative preferences and true well-being in fact involve. Our earlier discussions of the techniques that normative BLE deploys revealed, in detail, precisely how normative BLE relies on prior judgments of true well-being. But economists' technocratic expertise is limited to means and does not reach ends. (Once again, traditional law and economics was able to pursue its division-of-labor strategy—and deploy the rational actor model to sustain concrete policy outputs from modest substantive inputs—only because rational agents reveal what maximizes their welfare through their choices, and in this way fill the formal rationality deployed in traditional economic models with substantive content.<sup>204</sup>) Normative BLE scholars therefore have no reliable way to fill the void that descriptive BLE

---

204. And also, for the deontically inclined, choose in ways that realize their freedom and dignity.

has opened, which is why (as we have shown) they produce concrete recommendations only by making stipulative choices about precisely the questions of value that they purport to resolve, an approach that risks reproducing the biases—based on demography and professional disposition—that the scholars themselves bring to their work. In this way, when normative BLE occupies the field that descriptive BLE’s attack on rationality has opened up, the result is that the preferences of one social and demographic group—the group from which BLE scholars are disproportionately drawn—may dominate everyone else. The debiasing, nudges, and other interventions that technocrats favor risk being at best an imposition of the experts’ enduring biases on everyone else and at worst an effort to launder elite self-interest in a pseudo-scientific bath. This is precisely the state of play, we have argued, concerning the technocratically-waged campaign to assuage an expert-declared “retirement savings crisis.”

Finally, it is hard to see how the technocratic approach could yield any other result. If normative BLE wishes to preserve traditional law and economics’s division-of-labor strategy, even in the face of descriptive BLE’s attack on ordinary preferences and choices, it will eventually have to take up the full, interdisciplinary range of thought about what promotes human flourishing.<sup>205</sup> But while it is obviously sensible that policy should reflect all the values that it impacts, this is more nearly a tautology than a useful guide to action. No one can object to doing what is all-things-considered best, but no one can embrace this as a practical guide to *policymaking*. Moreover, where reasonable disagreement about what is best cannot be eliminated, even the most comprehensively informed expert opinion cannot possibly carry the authority that technocratic BLE aspires to. BLE therefore faces a dilemma: either it must cede its normative ambitions to other fields and content itself with purely descriptive analysis; or it must impose normative structure, even in the face of ordinary people’s contrary choices and stated preferences, risking the imposition of the BLE scholars’ notion of well-being.

We therefore seek an alternative to technocratic BLE that promises to integrate expert and ordinary opinion in productive ways, even accepting the skepticism about ordinary opinion that descriptive BLE so powerfully raises. We believe that democratic BLE is such an alternative.

### *B. Democratic Behavioral Law and Economics*

A second policymaking tradition integrates ordinary and expert opinion using a strategy that differs dramatically from the technocratic approach. Instead of enforcing a strict division of labor between ordinary and expert opinion, this democratic tradition encourages ongoing, discursive exchange between experts and ordinary people. We propose a new democratic behavioral law and

---

205. And also, for the deontically inclined, what respects human freedom and dignity.

economics that, consistent with this tradition, makes integrating expertise and ordinary people's views in an intentional way into one of its central ambitions.

Democratic approaches to policymaking understand experts as one part of a back-and-forth with ordinary people and, therefore, include both groups in discursive practices of reason-giving and justification. These traditions, which we now bring to BLE, have deep roots in legal theory and practice. Representative (as distinct from direct) democracy, for example, combines professional politicians who make policy and lead public opinion with mass elections that capture ordinary sentiment. Officials both shape and are shaped by popular movements, as professionals and citizens argue about how and why politics and policy should go.<sup>206</sup> Adjudication similarly combines the technical expertise of lawyers (judges and advocates), who frame the causes of action and legal theories that lawsuits pursue, with the lay sensibilities of parties who set the ends of litigation and juries who find facts. In this way, lawyers and ordinary people, as they interact, drive the development of common law precedent.<sup>207</sup> Finally, administrative procedure, especially notice-and-comment rulemaking, aspires to invite everyday citizens to explain their interests to regulators and to opine about the merits of proposed regulations. Citizens frame their concerns in ways that experts find compelling, and technical expertise is made accountable to public opinion.<sup>208</sup> The three practices, of course, cover discrete domains and differ in their details, but they all reject any rigid separation between experts and ordinary people and instead insist that both groups bring their sensibilities and capacities to a shared, deliberative practice of collective decision-making. While they vary in how successfully they achieve this aspiration, and while none succeeds entirely, the aspiration makes them all, in the broad sense of the term, democratic.

Democratic BLE joins and mimics these practices by pursuing a discursive integration of technical expertise and ordinary opinion, which directly and specifically addresses behavioral effects and the gap that they open up between revealed and normative preferences. We believe that the right response to the difficulties that descriptive BLE reveals is to abandon the division-of-labor strategy for integrating expert and ordinary opinion that economics has traditionally pursued in favor of promoting exchanges between technocratic experts and ordinary people, specifically designed to cure the defects of ordinary reasoning that BLE's descriptive work identifies and also the defects of paternalism that plague technocratic BLE. We call our approach democratic behavioral law and economics, and associate it with other democratic traditions in law and policymaking, because our approach uses behavioral insights to empower ordinary people, and because the path to empowerment runs through

---

206. See JAMES FISHKIN, *DEMOCRACY AND DELIBERATION: NEW DIRECTIONS FOR DEMOCRATIC REFORM* (1991).

207. See DANIEL MARKOVITS, *A MODERN LEGAL ETHICS: ADVERSARY ADVOCACY IN A DEMOCRATIC AGE* 171-211 (2008).

208. See JERRY MASHAW, *REASONED ADMINISTRATION AND DEMOCRATIC LEGITIMACY: HOW ADMINISTRATIVE LAW SUPPORTS DEMOCRATIC GOVERNMENT* 2-11 (2018).

deliberative engagements between expert technocrats and ordinary people. Building a comprehensive, fully elaborated democratic alternative to the technocratic traditions that dominate BLE as now practiced is the work of a field rather than of a single article. But we can light a path even if we cannot travel it; and we therefore conclude our argument here by outlining a theory of democratic BLE and applying that theory to the now familiar concrete case of retirement savings.

Unlike the technocratic approach, in which expert policymakers cure behavioral irrationalities paternalistically by manipulating ordinary people's choices, the democratic approach includes and indeed invites ordinary people into deliberation about policy. To do so, democratic BLE empowers ordinary people to step back from their naive choices, becoming self-conscious about the choices and the factors that led to them, and then to make the choices again from the new perspective of heightened self-consciousness. In particular, and crucially, our democratic approach aspires to empower ordinary people in a new way: not by lecturing them about what (experts say) they should rationally want, but by making them self-conscious of the biases that descriptive BLE identifies and the ways in which these biases influence everyday revealed preferences and choices.

We therefore embrace BLE's debiasing project but use a very different method from the one adopted by technocratic BLE. Where technocratic BLE asks experts to measure biases and purify preferences of the impacts of those biases, the democratic BLE that we propose draws those who will be regulated by behaviorally-inflected nudges into the process of identifying their own biases and designing the nudges themselves. Unlike traditional technocratic debiasing training, which seeks to purify actual preferences so that they match rational choice as "identified" by experts, this *deliberative debiasing* would involve educating people *about the insights of behavioral economics*, so that they came to share expert learning about their alleged irrationalities.<sup>209</sup> A businessperson who succumbs to optimism bias in assessing commercial opportunities, for example, would not be cured by having her bias technocratically measured and then being paternalistically manipulated so that her "choices" track preferences purified of the optimism. Instead, she would be invited to learn what descriptive BLE teaches about her optimism and its effects and then to assess new opportunities self-consciously, taking her now-acknowledged optimism bias into account. Our approach is novel, but far from unprecedented; we are inspired, for example, by George Loewenstein's and Emily Haisley's image of the economist

---

209. Philosophers loosely associate the technocratic approach to debiasing with Kant (on account of the connection between purified and categorical judgments) and the second with Hegel (and the dialectic method). This distinction is inspired by JONATHAN DANCY, *MORAL REASONS* 46, 147 (1993). We note that the first, Kantian, approach is best suited to cases in which there is no room for reasonable disagreement about what an unbiased judgment will yield and that the second, Hegelian, approach is best suited to cases in which reasonable disagreement is ineliminable.

as therapist.<sup>210</sup> We take this suggestion morally seriously, explaining how promoting self-consciousness about biases might save BLE from the normative trap that its descriptive successes have set.

When we call this approach “democratic,” we capture its deep structure rather than just giving it a label. The deliberative approach to debiasing, as compared to the purifying approach, fundamentally transforms the relationship between technocrats and ordinary people, between expertise and common choices. Under technocratic BLE, experts are paternalistic managers of ordinary behavior. But under democratic BLE, experts serve as tech-support, empowering ordinary people to make choices that they more fully and deeply embrace.

### *C. A Case Study of Democratic BLE*

All three contrasts—between the division of labor in technocratic BLE and the discursive engagement in democratic BLE, between purifying and deliberative debiasing, and between paternalism and tech-support—make concrete differences to how normative BLE built on the democratic model proceeds in practice. We revisit our case study of retirement savings to make our democratic approach to normative BLE concrete, display these differences between democratic and technocratic BLE, and shed further light on some of the theoretical arguments that we have already made.

Technocratic BLE uses expert-driven methods to purify preferences and identify the optimal savings rate and then asks how defaults might be used to induce otherwise recalcitrant workers to save at those rates (typically, to sustain 100 percent of their consumption on retirement). Even programs that pay lip-service to worker participation share in this basic structure. For example, traditional financial education does not invite workers to participate in determining what savings are optimal but instead begins from a stipulated optimal rate (fixed by expert opinion) and then seeks to teach workers how much they must save in order to achieve this rational result.<sup>211</sup>

By contrast, democratic BLE seeks precisely to engage workers in determining what savings rate is best and then makes policy responsive to what workers decide. It does this by replacing technocratically driven efforts to purify savings preferences with a deliberative process that invites workers to study actual savings behavior, to assess the forces that drive it, and to reflect on the mechanisms that govern savings among actual, imperfectly rational people. Whereas traditional technocratic financial education trains workers to know what

---

210. See George Loewenstein & Emily Haisley, *The Economist as Therapist: Methodological Ramifications of ‘Light’ Paternalism*, in *THE FOUNDATIONS OF POSITIVE AND NORMATIVE ECONOMICS: A HAND BOOK* 215 (Andrew Caplin & Andrew Schotter eds., 2011).

211. See, e.g., Paula Pant, *How Much Should I Save Each Month?*, TIAA (2022) <https://www.tiaa.org/public/learn/personal-finance-101/how-much-of-my-income-should-i-save-every-month> [<https://perma.cc/H8HP-MC27>] (recommending that “at least 20% of your income should go towards savings”).

(experts conclude) optimal savings requires and how to reach the optimum,<sup>212</sup> this democratic financial education aspires to bring workers to a new consciousness of their own values and behaviors. The workers themselves, now judging from a perspective that is aware of their potential biases, then say both what savings rate they think is best and how they think savings policy should be designed.

Fully elaborating democratic BLE, even for a single case such as retirement savings, requires resolving a thousand details of principle and practice. Here, we identify two buckets of questions—on how to achieve deliberative debiasing and on how policymakers should use those results—that a fully elaborated, practically implementable regimen must answer, and we begin (but only begin) to fill each.

First, a program of democratic BLE must explain how the deliberative debiasing that it proposes can be achieved. The financial education contemplated above would be time-consuming and expensive, of course—much more so than traditional financial education. Whereas traditional financial education can be stripped down to teaching two facts—the size of an optimal nest-egg and the savings rate needed to grow it—deliberative debiasing requires teaching a wide range of often difficult ideas. Workers would learn, for example, about the broad tendency to discount hyperbolically (of course, avoiding such jargon) and the effects that hyperbolic discounting has on current savings behavior and future regret. This is a challenging idea, and interpreting its intricacies in ways that can yield practical guidance requires engaging deep moral questions. Like every form of consciousness-raising, deliberative debiasing cannot be done on the cheap.

The education required for deliberative debiasing need not be universal, however, and its costs need not be incurred by, or for, all workers. Instead, deliberative debiasing concerning retirement savings might focus on a sample of workers, and their debiased judgments might then become inputs into savings policy, including through mechanisms that are already familiar to both democratic legislation and administrative law. BLE might borrow and adapt ideas concerning “deliberative polls” developed in the theory of deliberative democracy<sup>213</sup> to concentrate deliberative debiasing on representative samples of regulated people and then borrow from administrative procedure to give these representatives a role, alongside experts, in devising the nudges that BLE policy will deploy.<sup>214</sup>

---

212. See, e.g., Hunt Allcott & Dmitry Taubinsky, *Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market*, 105 AM. ECON. REV. 2501 (2015) (conducting such a short information intervention).

213. See generally FISHKIN, *supra* note 206; Bruce Ackerman & James Fishkin, *Deliberation Day*, 10 J. POL. PHILOS. 129 (2002); AMY GUTMANN & DENNIS THOMPSON, *WHY DELIBERATIVE DEMOCRACY?* (2009).

214. Poll participants might be *representative* in two senses of the term. First, and following the dominant tradition in deliberative polling, participants might be *statistically* representative of the general population. Alternatively, poll participants might be *politically* representative, in the sense of having been

Deliberative polls have now been used successfully in dozens of places around the world, showing significant knowledge gain and reconsideration of the views that participants entered with.<sup>215</sup> In such polls, a demographically and politically representative sample of participants is recruited from the relevant population.<sup>216</sup> The participants first receive briefing materials. Later, they have an intensive discussion with experts on various sides of a set of policy issues and also with one another. Having briefing materials and experts on various sides of an issue, along with discussions in which participants share their reasons and ask new questions of one another and of the experts, is important.<sup>217</sup> One of the reasons it is helpful to have a discussion with experts is that experts are unlikely to be able to anticipate all questions that the public might have, so discussion can bring out further issues. In addition, as James Fishkin notes, “[h]aving *competing* experts is important. Once [participants] realize[] that the experts disagree, it is much easier for participants to consider the competing arguments themselves rather than simply deferring to expert judgment.”<sup>218</sup> In the end, participants report their preferred policy options through voting by secret ballot to avoid “herding” to one answer.

Deliberative polls have been used successfully to handle even complex topics. For example, the Texas Public Utility Commission conducted a deliberative poll for its “Integrated Resource Planning,” which included questions like what mix of power the utility should use.<sup>219</sup> After listening to experts argue for and against increasing renewable energy in exchange for higher bills, the share supporting renewables rose from 52 percent to 84 percent.<sup>220</sup> The Commission adopted the results of the deliberative poll. Similarly, a deliberative poll of the citizens of Ulaanbaatar, Mongolia, resulted in a “participatory budget” adopted by the city. Now, by the National Law on Deliberative Polling, such procedures must be used more broadly for local development funds.<sup>221</sup> In Uganda, where—as in Mongolia—many citizens have few years of formal education, topics addressed by deliberative polls have included land-use zoning and transportation planning.<sup>222</sup>

Deliberative polls deployed in connection with democratic BLE would have a twist. Unlike existing deliberative polls, polls for democratic BLE would be aimed not at solving a question of collective policy choice, but rather at gaining self-awareness over one’s own potential behavioral biases.<sup>223</sup> And unlike

---

authorized by the general population (for example, through elections) to represent its perspective in the deliberative poll. We thank Joshua Teitelbaum for raising this possibility.

215. FISHKIN, *supra* note 43, at 74-75, 144.

216. *Id.* at 73-74.

217. *Id.* at 75.

218. *Id.* (emphasis added).

219. *Id.* at 160.

220. *Id.*

221. *Id.* at 99.

222. *Id.* at 106.

223. See Center for Deliberative Democracy, *supra* note 44. Note that there have been deliberative polls on behavioral issues. For example, Victoria had one on obesity. *Victoria’s Citizens’ Jury*

standard behavioral-economics information treatments, the briefing materials and experts deployed by the polls would focus not on the purifying project of “teaching” ordinary people what technocratic expertise reveals about the choices that they should make but rather on the deliberative project of empowering ordinary people to become self-aware about the mechanisms by which they, in fact, do choose. Deliberations, that is, would prominently include the insights generated by BLE’s descriptive project. In the context of retirement savings, as noted earlier, participants would learn about hyperbolic discounting—that people (presumably themselves included) may place an extra discount over all futures versus the present, over and above the discounts that they apply to farther over nearer futures, and that this pattern leads people simultaneously to want to save and to put off saving into the future. They would also discuss the normative frames that might be used to evaluate such behavior, including the forms of regret that hyperbolic discounting engenders. In both the discussion of the basic theory and evidence, as well as the discussion of the normative frames, the focus of deliberation would not be on persuading ordinary people to embrace what technocratic reason concludes are the “right” choices but rather on bringing ordinary people to self-consciousness about the patterns and mechanisms through which they in fact do choose. This is in keeping the democratic BLE’s commitment to rejecting the division-of-labor approach in favor of a discursive engagement between expert and ordinary opinion. By the end of the deliberative polling session, the ordinary people participating in the poll will have access to expert understandings of their own behavioral tendencies and will be in a position to combine these understandings with their own values and habits, to produce preferences that are debiased in the sense of coming from a perspective of informed self-consciousness.

It is always possible that the choice of briefing materials, or the design of the deliberations, will reintroduce familiar biases into deliberative polls or even introduce new ones. Perhaps expert opinion (and the biases that expert opinion can reflect) will dominate deliberations. Or perhaps differences of power and status from the world outside the poll will infect deliberations, so that dominant groups (races, genders, and so on) from the wider world will control discussions and impose their views on subordinated groups. But while both concerns are reasonable, well-designed deliberative polling mitigates each pitfall to achieve genuinely collaborative deliberation. As noted, expert presentations are selected to disagree with each other, and several decades of experience with deliberative

---

on *Obesity Insights Report 2016*, VICHEALTH (2016), <https://www.vichealth.vic.gov.au/-/media/ResourceCentre/PublicationsandResources/General/Victorias-Citizens-Jury-on-Obesity-Insights-Report-2016.pdf?la=en&hash=EEBF88790BDA90E3613B09FAB071D9C5EE927D65> [https://perma.cc/42J7-AQHR]. However, these jurors were prompted, “We have a problem with obesity. How can we make it easier to eat better?” thus presuming the problem and differentiating it from our approach. Also, the UK had deliberative consultation on savings in 2005-2006. *Security in Retirement: Towards a New Pensions System*, DEP’T FOR WORK AND PENSIONS (May 2006), [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/272299/6841.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/272299/6841.pdf) [https://perma.cc/SMY9-TWRB]. We cannot ascertain what participants were told, but we presume that the focus was not on individual behavioral failings because the stated focus was “the tough pensions choices we face” as a society.

polling has produced best practices concerning how to use competing experts to ward off domineering expertise.<sup>224</sup> Moreover, expert presentations that are focused on reasoning itself—on teaching about behavioral effects—rather than on which outcomes are best, tend naturally to empower rather than to dominate ordinary reasoning. Finally, deliberative pollsters know how to insulate the deliberations that they facilitate from hierarchy and subordination in the outside world, and empirical study shows that in a set of well-designed deliberative polls, “if one looks at the initial positions of the more advantaged (the more educated, the males, the rich, etc.) on policy indices on the issues deliberated about, there is no pattern of movement . . . in the direction of those positions. The more advantaged are not dominating the discussion by imposing their views.”<sup>225</sup>

Similarly, though one might be concerned about challenges recruiting certain samples of the public (especially the rich or the poor), these polls have been able to achieve demographic representativeness.<sup>226</sup> This makes sense, given that a minuscule share of the population must participate in the survey. More broadly, the effectiveness and accuracy of modern sampling techniques entail that truly representative polls can be constructed and administered using sample sizes small enough to be cost-effective, certainly when set against the enormous amounts at stake in the policy areas that democratic BLE might be deployed to address.

Of course, we cannot claim that any good-faith method can achieve a “neutral” description of evidence. Our claim is that it is often better to have the people themselves contributing to drawing the policy implications from the behavioral evidence, based in part on experts’ disagreeing with one another, rather than subsets of potentially biased experts deciding. Nor can we solve concerns about interest groups in the policymaking process, which would likewise also be present with technocratic BLE. Nor can we rule out the possibility that bad-faith decision-makers—who could cherry-pick experts—might manipulate a deliberative poll. We are merely trying to improve the information that policymakers have at hand to reach their decisions.

This approach is novel, to be sure, but it is also practicable and effective. Suppose that policymakers in a state—say, Michigan—that has not implemented an auto-IRA are considering whether to adopt one and, in case they do, how to design it.<sup>227</sup> Recall that this is a live policy issue: auto-IRAs are currently spreading across the states.<sup>228</sup> Suppose further that administrators, unwilling to

---

224. *Id.* at 75, 77.

225. *Id.*

226. *Id.* at 94, 118, 135 n107, 157, 191, 197, 215 n57.

227. Note that although we focus on government policy, analogous deliberations could occur in many settings. For example, they need not even involve the government. Recall that employers are adopting paternalistic choice architectures that nudge employees into enrollment. An employer considering how to design the choice architecture of retirement savings plans could convene a group of its workers and ask a similar set of questions.

228. For example, Colorado just implemented one of these programs. Ellen Stone & Brian Kearney, *Colorado Enacts State-Run Auto-IRA Program*, *MERCER* (Aug. 27, 2020),

impose (even through nudges) expert biases on ordinary people, want to know what the people themselves think after having a chance to deliberate. Rather than looking to technocratic BLE to purify savings preferences of their biases and design a choice architecture that induces workers to “choose” to save as if they were rational, the administrators might convene a deliberative poll.

Michigan would first recruit approximately 500 participants from across the state (a sample small enough to be feasible and cost-effective, but large enough to be representative<sup>229</sup>). They could meet in-person in Lansing or by videoconference. Travel expenses would be paid, along with a small honorarium. The participants would be representative of Michiganders along the dimensions of sex, race, education level, income, political affiliation, and the like. The population from which the poll-participants were drawn would be determined with the authority that the poll’s results aspire to in mind. In the case of private pensions, plan policies must answer to the workers whose savings they influence. The relevant population in this case is therefore anyone who could work or has worked, so essentially all adults.<sup>230</sup> In other cases, poll participants might be selected differently, by drawing from narrower groups—for example, in polls designed to inform savings policy for a particular industry.

As explained, the poll would aim to bring participants to self-consciousness about the behavioral effects (identified through descriptive BLE) that might lead their revealed preferences regarding savings to depart from their normative preferences. Participants would have information sent to them beforehand: on current savings rates across the income distribution, how auto-IRA programs work, preliminary and simplified discussions of potential biases that savers could have, and evidence on the significant impacts of defaults. They would also get an interactive simulator that shows how defaults impact savings rates and how savings rates impact consumption over the course of one’s lifetime, particularly for low-income individuals, who risk harm to their everyday needs from saving but also may suffer at retirement from too little saving. Even before meeting, participants would be asked their opinions on retirement savings, to begin the engagement.

The participants would then meet for a day. There would be expert presentations. One expert would explain the kind of biases that may reduce savings too much, such as hyperbolic discounting. Another expert would argue against the idea that hyperbolic discounting represents a problem that leads to savings rates that are too low. Experts would also argue for and against the auto-IRA program specifically, pointing alternately to under-saving but also to the needs of day-to-day life, especially for low-income families. Participants would

---

<https://www.mercer.com/our-thinking/law-and-policy-group/colorado-enacts-state-run-auto-ira-program.html> [https://perma.cc/4DDS-MHJ3].

229. FISHKIN, *supra* note 43, at 74, 80.

230. We note that we would include retirees in this case to provide perspective for the later lives of current workers. Recall that our goal here is not informing individuals to get them to change their behavior, but rather helping a small group reflect on what they think policy should be. And having something like their future selves as part of the conversation is helpful for that.

then break up into small groups of a dozen or so, allowing them to discuss these issues with each other, young and old alike learning from each other's different vantage points. Experts would be on-hand to answer questions. There would again be time for individuals to use an app that simulates the effect on savings at retirement of various plans—with and without auto-IRAs and with various income thresholds for exemption—as well as various savings rates.

Finally, this learning would be directed towards the problem at hand: deciding what savings rates are optimal and what choice architecture should govern retirement savings.<sup>231</sup> Poll participants, having learned about biases that previously potentially influenced their savings behavior would be asked how much they will save going forward (now that they know what they know),<sup>232</sup> whether auto-IRAs should be adopted for workers in general, and (if so) whether low-income individuals likely to need their current funds the most should be exempt from the program. Responses would be broken out by demographic and other categories, including in particular by income level, to gain insight into both what the overall population thinks and also what lower-income individuals think.

The second bucket of questions to frame democratic BLE arises once deliberative polls have been completed. In their broadest form, these questions ask how the results of deliberative polling should be deployed to give deliberatively debiased ordinary preferences an influence over policy. Democratic BLE must develop principles to say both where in the policymaking process the results of deliberative behavioral polling should come into play and, wherever they do come into play, what authority deliberative polls should have. We once again make no pretense to offer complete or final answers to these questions but content ourselves with some partial and tentative suggestions.

To begin with, the results of deliberative polls might enter the policymaking process in any number of ways, at any number of stages. Where a choice architecture leaves room for ordinary people to make decisions, they might simply be informed of the results of deliberative polling. For example, workers choosing how much to save (either through opt-in or opt-out savings plans) might be told how much poll participants (both in general, and those that best represent the workers in question) chose to save and how participating in the poll changed participants' choices. In other circumstances, where policymakers decide which way and how firmly to nudge or adopt mandatory rules, deliberative polling can inform these decisions also. For example, if poll participants overwhelmingly

---

231. Of course, if one really has “multiple selves” over time and there is a question of justice between the selves, then this method will be limited—like technocratic BLE—because it only knows the views of the present self. Nevertheless, democratic BLE still expands the consciousness of the citizen, moving closer towards an understanding of one's whole life.

232. One could also combine the virtues of deliberative polling with the benefits of seeing actual behavior, in which participants have real stakes. Some questions are about policy and can only be asked in the abstract. But, for the savings rate into a savings plan, the participants could actually choose whether to enroll in a savings plan and then, if they enroll, set a savings rate. Our setup could observe these actual choices, where participants have skin in the game. The reason that we can do this, whereas most deliberative polls cannot, is that our proposal is primarily targeted at individuals' behavior, not reconciling preferences across adverse interests.

reject auto-IRA savings plans, even after having learned of the biases that cause them to save “too little,” then this cuts against adopting such plans. More nuanced uses of deliberative polling results are also possible. If, for example, deliberative debiasing had the smallest impact on the savings preferences of the lowest-paid workers and they were mostly against auto-IRAs, then this would suggest that these workers’ choices not to save reflect current budget constraints (which are distinctive to them) rather than short-sightedness.

Furthermore, the results of deliberative polling might exert a smaller or greater influence at each of these policy sites. At the weak end of the continuum, deliberative polls might serve as no more than a reality-check, or a fresh perspective, on experts engaged in the purifying forms of debiasing that now dominate technocratic BLE. The tremendous stakes of the policy decisions on which deliberative polls can shed light, along with the comparatively small cost of conducting a deliberative poll, argue for at least using them for this purpose. For example, if workers prefer to save at current levels and to not be defaulted into IRAs, that should give even experts who continue to embrace technocratic BLE pause before purifying revealed preferences in ways that declare that normative preferences require materially greater savings. At the strong end of the continuum, the results of deliberative polling might be decisive for policymakers, subject to other requirements like statutory mandates. Once again, if substantial majorities of poll participants re-affirmed their prior behavior even from the behaviorally self-conscious perspective achieved through the poll, then this might flatly reject any policy changes that nudged (let alone mandated) ordinary people to depart from their revealed preferences.

A middle ground that might be used in notice-and-comment rulemaking by agencies is also possible. Rulemakers might not be bound by the results of the deliberative poll, but they might have a duty to explain in their rulemaking why they differ if they adopt a rule counter to what the majority or supermajority of the poll chooses, forcing rulemakers to confront the results. The adequacy of that reasoning for rejecting the results could be judicially reviewable, much as federal administrative agency explanations of how their decisions consider relevant issues raised by experts are today judicially reviewable.<sup>233</sup>

Note that none of these approaches embraces blunt rule by plebiscite, where manipulation and ignorance can easily dominate. Rather, as proponents of “deliberative democracy” argue, in a democracy we care about the will of the people—and surely it is the *informed* will of the people that matters most.<sup>234</sup> We propose ways in which BLE’s descriptive insights might be drafted into the project of constructing and implementing an informed popular will.

Finally, it is important to emphasize the scope of what we are proposing here along four dimensions. First, again, we are by no means arguing that

---

233. 5 U.S.C. § 553 (2018) (requiring “consideration of the relevant matter presented” to the agency; *Motor Veh. Mfrs. Ass’n v. State Farm Ins.*, 463 U.S. 29 (1983) (describing courts’ “hard look” review of agency rulemaking).

234. *Id.* at 72.

standard technocratic economic evidence be discarded. For all its problems, we embrace this evidence. It usefully advances knowledge, and deliberative polling has its own set of downsides, including the challenges inherent in informing the public—even at length, with experts—about the behavioral evidence. But, given the huge stakes of behavioral interventions and the relatively modest expense of a deliberative poll, we believe that it is often well worth the cost to consult well-informed, representative sets of the public about what they actually want when they know their own biases, at least as a supplement to traditional expertise.

Second, we do not propose that our deliberative polls, and the empowerment that they produce, will sustain permanent changes in the behaviors of people who participate in them. We propose only that participation will lead newly self-conscious participants to prefer different policies in the short run; and we argue that the distinctive preferences of empowered poll participants provide useful information for policymakers about which policies in fact promote the well-being and authentic freedom of imperfectly rational agents.

Third, while we acknowledge that the cost of deliberative polling must be taken into account in deciding whether to adopt our method in any given case, we also note that the scope of many government programs is so large that it is worth incurring significant costs to improve them. Take the example of the Michigan retirement program. Michigan's population is about 10 million people. Earlier studies using deliberative polls have provided an honorarium of \$100 for a weekend day, plus travel and lodging expenses.<sup>235</sup> Suppose that these payments, together with administrative expenses and expert fees, add up to \$500 per person, so that a deliberative poll of 500 participants would cost about \$250,000. Suppose also that 1 million Michiganders would benefit for a decade from the improved decision-making that the poll produced. How much would each have to benefit to make the poll worth its price? Only a miniscule 2.5 cents per person per year.<sup>236</sup> This lesson—which follows the logic of sampling—generalizes. Whenever a policy effects large numbers of people, only extreme skepticism about the expected benefits of democratic BLE might justify rejecting our approach on cost grounds.<sup>237</sup>

Fourth, we are not proposing that all questions of policy respond to deliberative polling. Rather, we are suggesting that a certain subset of policies—those whose proper aim is promoting experienced well-being, in contexts where behavioral effects open a gap between revealed and normative preferences—should often embrace deliberative polling on the model of democratic BLE. We do not take a stand either way on other types of questions, which are beyond our scope. We therefore do not propose democratic BLE for policy areas in which

---

235. FISHKIN, *supra* note 43, at 80, 128 n.18.

236. We note that, of course, one could discount future benefits, which would push up the needed benefits a little.

237. Other potential costs of deliberative polling are delay and even legal action, as another procedure to follow. While we believe that these costs should be taken into account, we are hopeful that appropriate institution design could leave such costs small. Conversely, one state can learn from a poll in another state, increasing the value of the poll.

rights, fairness, or conflicts among different interests are dominant concerns (for example, cases involving basic liberties, equal citizenship, or the regulation of externalities). Nor are we proposing our approach for areas with hardened partisan views, where discussion may have less of an impact. Nor are we talking about areas where revealed preferences over outcomes reliably track well-being, but expertise is required to determine which means best promote the reliably preferred outcomes (for example, in medicine, where patients reliably prefer to be cured of disease but depend on experts to say which medicines and treatments are safe and effective). Instead, we propose to deploy democratic BLE in cases, such as retirement savings, where normative preferences and experienced well-being determine good policy, but behavioral effects make revealed preferences unreliable guides to optimal conduct and policy.

#### *D. Democratic BLE in Context*

In these circumstances, deliberative polling, with a heavy focus on bringing ordinary choosers to self-consciousness about BLE’s descriptive insights, promises to complement the technocratic approach that dominates BLE policymaking today with a democratic behavioral law and economics—“democratic” not in the narrow sense of voting and elections but in the broad sense captured by Robert Dahl when he defined democracy as “the continuing responsiveness of the government to the preferences of its citizens, considered as political equals.”<sup>238</sup> A democratic turn is apt given the difficult normative problems that BLE-inflected policy must address. We often turn to democratic deliberation when called on to make collective choices in the face of ineliminable moral uncertainty.

Indeed, deliberative democracy naturally fills the role in behavioral law and economics that the market plays in traditional law and economics. When individual agents are free and rational, it can make sense (in certain circumstances) to leave them to their own devices, allowing private choices to produce efficient and just outcomes within a framework of market exchange established and policed by the state. But when agents are not free and rational, then there is no good substitute for collective decision. Technocratic BLE makes such decisions hierarchically, based on the presumptions of experts. By contrast, democratic law and economics proposes to take such collective decisions using procedures structured to include participation from all affected parties, treated as equals. In this way, democratic law and economics brings the egalitarian sensibilities that support decentralized markets—that no person’s values carry greater natural warrant or authority than any other’s and, as Hayek warned, that economic planning therefore tends towards tyranny<sup>239</sup>—to problems that

---

238. ROBERT DAHL, *POLYARCHY: PARTICIPATION AND OPPOSITION* 1 (1971).

239. *See generally*, FRIEDRICH HAYEK, *THE ROAD TO SERFDOM* (1944).

(because individual agents are irrational) require centralized coordination.<sup>240</sup> Much as mass democracy deploys joint deliberation to align the practical imperative of collective political decision with individual freedom and dignity,<sup>241</sup> so democratic BLE deploys deliberative debiasing to align individual dignity with the practical imperatives of managerial administration.

These remarks do not so much elaborate a policy program as announce a research agenda. Democratic BLE must study how to deliver the training that deliberative debiasing requires—by what means, in what settings, and to which people. For example, it must study how to integrate this democratic approach into the policy advice that Executive Order 13,707 directs the government “nudge unit”—at present populated exclusively by experts—to provide to administrative agencies, which are in turn directed to apply these insights.<sup>242</sup> Democratic BLE might even study how to present the results of democratically driven deliberative debiasing to non-experts, in order vicariously to empower ordinary people who have not participated in deliberative polls.

The path that democratic BLE must traverse in answering these questions is long; but it leads, morally, forward.

## Conclusion

The descriptive and normative facets of traditional law and economics work together: the normative project validates the patterns that the descriptive project predicts; and the descriptive project saves the normative project from having to make substantive judgments about what ends are worth pursuing. To be sure, the most extravagant version of this argument—in which traditional economic thought claimed to be both normative and at the same time objective and indeed purely technocratic<sup>243</sup>—was never persuasive.<sup>244</sup> But more modest constructions, which emphasize the utilitarian and liberal bases of consumer sovereignty in economic markets, at least worked on their own terms.

The descriptive and normative projects converge on a single, familiar, legal regime. Laws and policies that inform rather than command leverage agents’

---

240. Of course, no objective barometer of “well-being” exists, which is precisely the point we argue. So, we cannot claim that our method best achieves well-being. But we can say that, through engagement between experts and the people themselves—and then letting the people’s choices be the primary output—democratic BLE aims to develop as well as possible what the people’s own sense of their reflected-upon preferences are, treat those preferences with respect, and implement policy on their basis.

241. This is captured most powerfully in Rousseau’s suggestion that democratic political engagements can enable the general will to arise out of the will of all. *See generally*, JEAN-JACQUES ROUSSEAU, ON THE SOCIAL CONTRACT (1762).

242. Exec. Order No. 13,707, 3 C.F.R. § 13707 (2015). *See also* Adam Smith, *Utilizing Behavioral Insights (Without Romance): An Inquiry into the Choice Architecture of Public Decision-Making*, 82 MO. L. REV. 737 (2017) (noting that this executive order “puts behavioral insights on the same footing as cost-benefit analysis in guiding regulatory policy”).

243. *See, e.g.*, N. GREGORY MANKIW, PRINCIPLES OF ECONOMICS 145 (7th ed. 2015); Robert Heilbroner, *Economics as a “Value-Free” Science*, 40 SOC. RSCH. 129 (1973).

244. *See, e.g.*, THE END OF VALUE-FREE ECONOMICS (Hilary Putnam & Vivian Walsh eds., 2012); ELIZABETH ANDERSON, VALUE IN ETHICS AND ECONOMICS (1995).

rationality and respect their liberty, and in this way produce optimal and just outcomes. Command and public coercion should be deployed only where circumstances render private choices no longer rationalizable—where private power, externalities, and other market failures break the link between choice on the one hand and welfare and freedom on the other. Reasonable people—both left and right—might reject this moral and political vision. But notwithstanding powerful critiques, the tradition has endured.

BLE arises against this backdrop to reject the model of the rational actor that both drives and unifies the traditional approach. BLE scholars insist that economic agents do not, in fact, optimize in the fashion that rationality requires. Instead, they deploy a wide range of heuristics and display a wide range of biases that simultaneously break the connection between revealed preferences and experienced well-being and bend agents' choices to the frames in which they are made.

BLE presents itself as a program of modest reform.<sup>245</sup> And the behavioral turn does not in fact pose any fundamental threat to traditional law and economics's descriptive agenda. The biases that BLE scholars identify may be irrational, but they are not for this reason arbitrary, capricious, or otherwise unpredictable. To the contrary, these biases take on a few prominent and tractable forms. By modelling the biases, BLE promises to predict behavior.

This approach is perfectly consistent with economics understood as a descriptive science. "There is," Kenneth Arrow has explained, "no general principle that prevents the creation of an economic theory based on other hypotheses than that of rationality."<sup>246</sup> In fact, "any coherent theory of reactions to stimuli appropriate in an economic context . . . could in principle lead to a theory of the economy."<sup>247</sup> Indeed, insofar as people are in fact *not* rational, behavioral approaches—which describe people not as they ideally would be but as they actually are—can *improve* on traditional law and economics, as a descriptive project. This is what led Jolls, Sunstein, and Thaler to celebrate that BLE "offers the potential to be law and economics with a higher 'R<sup>2</sup>'—that is, with greater power to explain observed data."<sup>248</sup> A common—indeed dominant—view therefore treats behavioral approaches as reformist: as proposing a friendly amendment to traditional law and economics.

---

245. See, e.g., Roberto Weber & Colin Camerer, "Behavioral Experiments" in *Economics*, 9 EXPERIMENTAL ECON. 187, 187 (2006) (noting that "most behavioral economists have the goal, not of developing an alternative to economic theory and methods, but instead to incorporate new assumptions and methods into mainstream economics research"); Tanina Rostain, *Educating Homo Economicus: Cautionary Notes on the New Behavioral Law and Economics Movement*, 34 L. & SOC. REV. 973, 976 (2000) (explaining that "[a]dvocates of behavioral law and economics . . . must incorporate a modest conception of social science"); Thomas Ulen, *The Growing Pains of Behavioral Law and Economics*, 51 VAND. L. REV. 1747, 1748 (1998) ("Behavioral law and economics does not attempt to undo any of the remarkable accomplishments of law and economics. Rather, it is an attempt to refine.").

246. Kenneth Arrow, *Rationality of Self and Others in an Economic System*, in RATIONAL CHOICE: THE CONTRAST BETWEEN ECONOMICS AND PSYCHOLOGY 201, 202 (Robin Hogarth & Melvin Reder eds., 1987).

247. *Id.*

248. Jolls, Sunstein & Thaler, *supra* note 26.

BLE stands in a very different relationship to the normative agenda of traditional law and economics, however. In respect of values, BLE proposes nothing short of a revolution. The behavioral revolution, moreover, destroys the incumbent normative order without erecting a new one in its stead. The descriptive program of BLE pulls the rug out from under traditional law and economics's normative program, and BLE contains no program of its own that might serve as a replacement. In this way, BLE threatens to shake law and economics, writ large, to its core.

If agents are not rational—but rather subject to pervasive manipulation—then their choices are no longer normatively self-authorizing. Being erratic rather than coherent, choices do not reliably promote well-being. And being manipulable rather than autonomous, choices cannot claim the presumptive authority associated with rational choice or human freedom and dignity. Traditional law and economics drew on deep wells of value theory concerning welfare and freedom. The behavioral revolution drains both wells dry.

The behavioral revolution therefore breaks the unity that held traditional law and economics together. The descriptive account of human action that BLE develops—at least insofar as it departs materially from the rational actor model—no longer supports the normative vision of welfare and freedom that traditional economics held and erects no new ideal in its place. Rather, BLE requires independent ideas—drawn from outside of economics—to provide its norms. These difficulties pose no threat to BLE's descriptive ambitions. But BLE's normative ambitions are a different matter: whereas the rational actor model gave traditional law and economics not just a descriptive method but normative foundations as well, BLE, as a normative enterprise, stands at the mercy of other disciplines.

Moreover, such normative vertigo is the inexorable and indeed immediate consequence of the behavioral turn. The model of the rational agent allowed traditional law and economics to forge a link between facts and norms. Abandoning the model—no matter how well-justified based on descriptive findings—leaves BLE normatively untethered. It is a case of the old saying that while it is true that one cannot make an omelet without breaking any eggs, it is also true that one can break a great many eggs without making an omelet.<sup>249</sup>

The very language that BLE scholars deploy betrays the problem. BLE scholars use the term “normative preferences” to identify the alternatives that will maximize well-being, independent of which alternatives are chosen in fact. But this use of the word “preferences” obscures the normative structure of BLE and invokes forms of argument that BLE's descriptive project rejects. Preferences properly-so-called simply are comparative attitudes that govern choice.<sup>250</sup> They are connected to well-being, and therefore to normativity, only contingently, through the rationality assumptions that traditional law and

---

249. Hannah Arendt, *The Eggs Speak Up*, *ESSAYS IN UNDERSTANDING: 1930-1954*, at 270, 275 (Jerome Kohn ed., 1993).

250. MAS-COLELL ET AL., *supra* note 22.

economics deploys and that BLE's descriptive project rejects. When BLE scholars, having abandoned rationality, define "normative preferences" directly in terms of well-being, they necessarily turn away from choice and therefore also from preference as a basis for normative insight. This is fine, so long as BLE provides alternative normative foundations. But BLE provides none, and the phrase "normative preferences" obscures the difficulty and suggests that the BLE's norms enjoy choice-based foundations that BLE's descriptive project denies. Leading practitioners thus often fail to acknowledge just how deep into the normative foundations of law and economics their descriptive results reach. BLE remains in denial of its revolutionary character.

The charisma of economics arises in large part from its claim to know what is right. But behavioral findings shake that method to its core. All that remains is an attempt to suss out well-being in different ways. But, while the empirical and theoretical tools of economics remain useful, the inherent uncertainties of knowing what the good life is require more than an economics training can provide. These uncertainties raise age-old questions of philosophy, theology, and even literature and art. BLE describes itself as a largely technocratic enterprise when, in fact, it can be no such thing. Of course, BLE can contribute expertise, but economics can also contribute expertise to the provision of medical care; that does not mean that economists should be our doctors the way that economists are in the policymaking driver's seat on a range of behavioral topics, through both scholarship and government behavioral insights teams spreading around the world. Doing so risks letting the peculiar biases of these experts direct policy.

While we applaud the descriptive work of BLE, as well as many of the policy suggestions generated by the normative work of BLE, ultimately any suggestion that these experts can define the policy that maximizes well-being or respects human dignity is illusory. Instead, law and economics—and BLE in particular—should own up to its inadequacies and try something new.

These critiques naturally lead to our proposals for reform. We develop a new approach, democratic BLE, which deploys behavioral insights and participatory practices to draw on the informed decisions of the people themselves. We seek to empower ordinary people to overcome the behavioral biases that—descriptive BLE so persuasively shows—otherwise prevent observed choices from maximizing well-being or achieving authentic freedom.

We propose a specific and novel method for empowerment, which is informed by our diagnosis of the normative defects that plague BLE, in its conventional technocratic form. Conventional responses to the biases that descriptive behavioral economics identifies seek to bring behavior directly into alignment with what rationality, as determined by experts, requires. Whatever their intramural differences, commands, nudges, and even conventional education—which instructs ordinary people about their true interests—all share this basic technocratic bent. These approaches all debias ordinary people by purifying their preferences of irrationalities. But because economists have no special insight or even training concerning human flourishing, they have no authority to determine what purified preferences would be. And insofar as

economists are unrepresentative, and themselves hold unusual views about what is good for people, the technocratic approach risks merely replacing ordinary people's biases with the economists' own.

We therefore propose to debias in a new deliberative way, by empowering ordinary people. We propose to use deliberative polls to allow ordinary people to overcome their biases and then to use the outputs of deliberative polls to inform policy. The deliberative polls that we propose would not lecture participants about what (experts say) they should rationally want—again, this would simply exchange one bias for another. Instead, our deliberative polls would focus on teaching BLE's descriptive insights, to make participants self-conscious about their own biases and the ways in which these biases influence their naive preferences and everyday choices. Where standard information treatments emphasize knowledge-transfer, our approach emphasizes consciousness raising; where standard information treatments are didactic, our approach is therapeutic; where standard information treatments emphasize instruction, our approach emphasizes empowerment.

This emphasis on empowerment returns democratic BLE to traditional law and economics's normative roots. Traditional law and economics, recall, could sustain strong normative conclusions about efficiency and freedom even without substantive normative inputs because it could trust ordinary people, trading in markets, to decide about values for themselves. This attitude—a respect for ordinary people—gave traditional law and economics a sort of democratic sensibility. The behavioral revolution's descriptive achievements—precisely on account of their power—demoted ordinary people. Technocratic BLE turns to economists' own values to fill the normative gap, but economists lack the expertise or authority to legitimate this approach. Democratic BLE, by contrast, empowers ordinary people to overcome their own biases and, working with expert helpers, to reassert control over their own lives.