



TOURO UNIVERSITY
JACOB D. FUCHSBERG LAW CENTER
Where Knowledge and Values Meet

**Digital Commons @ Touro Law
Center**

Scholarly Works

Faculty Scholarship

2020

Individual Preferences in Policy Analysis: A Normative Framework

Gabriel Weil

Touro Law Center, gweil2@tourolaw.edu

Follow this and additional works at: <https://digitalcommons.tourolaw.edu/scholarlyworks>



Part of the [Administrative Law Commons](#), [Civil Law Commons](#), [Constitutional Law Commons](#), and the [Jurisprudence Commons](#)

Recommended Citation

Gabriel Weil, Individual Preferences in Policy Analysis: A Normative Framework, 50 *Tex. Env't. L.J.* 55 (2020).

This Article is brought to you for free and open access by the Faculty Scholarship at Digital Commons @ Touro Law Center. It has been accepted for inclusion in Scholarly Works by an authorized administrator of Digital Commons @ Touro Law Center. For more information, please contact lross@tourolaw.edu.

INDIVIDUAL PREFERENCES IN POLICY ANALYSIS: A NORMATIVE FRAMEWORK

BY GABRIEL A. WEIL

I.	Executive Summary	55
II.	Introduction	56
III.	Two CBA Controversies Connected by Time Preference	57
IV.	Discounting the Future	59
	A. A Parfitian Defense	60
	B. Opportunity Cost	62
	C. Policy Regime Coherence	65
V.	Valuing Life	70
VI.	Well-Being Analysis	78
	A. Mortality Risk in WBA	78
	B. Time Discounting in WBA	83
VII.	Happiness vs. Preference Satisfaction	87
VIII.	Laundered Preferences	94
IX.	The <i>Nudge</i> Debate	96
X.	Other Alternatives to CBA	99
XI.	Conclusion	103

I. EXECUTIVE SUMMARY

Measures of individual preferences are a key input in cost-benefit analysis. However, behavioral science has raised questions about the rationality of these preferences. The *Nudge* thesis relies on that scientific research to prescribe interventions to influence individual choices. However, the more modest step of limiting reliance on these preferences in evaluating non-paternalistic government policies has not been taken up. We lack a consistent theory of when public policy should defer to these preferences, with legal and policy advocates adopting ad hoc result-oriented approaches. I argue that policymakers should be prepared to override individual preferences in cases where their only plausible rational justification(s) sever their connection to social welfare, undermining their normative motivation. For time discounting, this means eliminating the pure time preference component of the discount rate for most purposes. For valuing mortality risks, it implies shifting from the value of a statistical life method to a modified value of a statistical life-year method.

II. INTRODUCTION

Measures of individual preferences are a key input for cost-benefit analysis (CBA), the dominant mode of formal policy analysis in the United States.¹ However, behavioral science is increasingly demonstrating that individual preferences are often inconsistent, unstable, subject to systematic biases and framing effects, difficult to determine, and arguably irrational. For instance, willingness to pay (WTP) and willingness to accept for a lottery are weakly correlated at best,² an individual's preference between two options can flip based on whether they are evaluated separately or together,³ and people report higher WTP for a given public good when they are told it is more costly to provide.⁴ This research inspired Cass Sunstein and Richard Thaler's *Nudge* thesis—also known as libertarian paternalism—that prescribes crafting choice architecture to influence individual choices in normatively appealing directions consistent with enlightened preferences.⁵ It has not, however, significantly altered the reliance of regulatory bodies on measures of individual preferences in conducting CBA and setting policy.⁶

This article will argue that policymakers should not defer to individual preferences in cases where the only plausible rational justification(s) for them lack normative motivation, regardless of the merits of nudges. In particular, the component of the discount rate attributable to pure time preference should be eliminated because the only plausible rationale for pure time preference—Parfitian ideas about attenuated connectedness to one's future self—severs the connection with social welfare that justifies CBA. Likewise, policymakers should shift from the value of a statistical life (VSL) method to a modified value of a statistical life-year (VSLY) method. Arguments for retaining the VSL based on preferences fail. To the extent the claimed pattern of preferences holds, it still cannot be defended in normative terms.

The remainder of this article proceeds as follows. Part II articulates the tension in both pro- and anti-regulatory advocates' positions on discount rates and valuing life, elucidating how they embody inconsistent positions regarding deference to individual preferences. Part III examines the role of pure time preference in discounting, concluding that it is not a proper component of the discount rate, for most purposes. In doing so, it considers three defenses of pure time preference: (a) Parfitian ideas about attenuated connectedness with future selves, (b) opportunity cost, and (c) policy regime coherence.

1 MATTHEW ADLER & ERIC POSNER, *NEW FOUNDATION OF COST-BENEFIT ANALYSIS* (2006).

2 Colin Camerer, Jonathan Chapman, Mark Dean, Pietro Ortoleva, & Erik Snowberg, *Willingness to Pay and Willingness to Accept are Probably Less Correlated Than You Think* (Nat'l Bureau of Econ. Res., Working Paper No. 23954, 2017).

3 Cass R. Sunstein, *On Preferring A to B, While Also Preferring B to A*, 30(3) *RATIONALITY AND SOC'Y* 305 (2018).

4 J. Baron & N.P. Maxwell, *Cost of Public Goods Affects Willingness to Pay for Them*, 9 *J. BEHAV. DECISION MAKING* 173 (1994).

5 CASS R. SUNSTEIN & RICHARD H. THALER, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* (2008).

6 Sunstein acknowledges the possibility that behavioral market failures justify adjustments to WTP and corresponding VSL estimate, writing that "further conceptual and empirical work needs to be done on these issues." CASS R. SUNSTEIN, *THE COST-BENEFIT REVOLUTION* 55 (2018).

Extracting the component of market rates of return that reflects pure time preference results in a lower discount rate, meaning the future is valued more highly.

Part IV discusses mortality risk, applying similar reasoning as in Part III to conclude that a modified VS LY approach should replace the VSL method. Part V considers well-being analysis (WBA) as an alternative to CBA, concluding that it confronts difficulties in its treatment of discounting and mortality risk that are comparable to those facing CBA. Part VI evaluates whether social welfare is best conceived of in terms of preference satisfaction or subjective well-being. Part VII discusses Matthew Adler and Eric Posner's concept of laundered preferences and how their framework could accommodate the account of preferences articulated in the foregoing parts.⁷ Part VIII analyzes some key objections to the *Nudge* thesis, arguing that even if these objections hold, they do not justify deference to the relevant preferences when evaluating non-paternalistic regulations. Part IX considers feasibility standards, safety maximization standards, cost-effectiveness analysis, and other alternatives to CBA, arguing that they generally lack normative motivation and that CBA is preferable for most major rules. Part X concludes.

III. TWO CBA CONTROVERSIES CONNECTED BY TIME PREFERENCE

Consider two related controversies in CBA: the value of a life and the social discount rate. The choice between the VSL method and the VS LY method is important for several environmental, health, and safety regulations. Defenders of retaining the VSL method point out that direct revealed and stated preferences studies of older people do not reveal a substantially lower WTP to avoid near-term mortality risks, so discounting the value of their lives with VS LY measures would violate their preferences.⁸ They would prefer to treat accidental death at age sixty as equivalently bad as death by the same mode at age thirty, even though the thirty-year-old can expect to live about twenty-six more years than the sixty-year-old.⁹ Doing so implicitly embraces a high value for the rate of pure time preference, a component of the discount rate that reflects the lower value people place on an otherwise identical (including in probability) cost or benefit that is delayed in time.

When it comes to the social discount rate, the alignment of interests flips such that anti-regulatory advocates and scholars tend to support use of a high social discount rate, consistent with the high rate of pure time preference most people exhibit in their market behavior.¹⁰ A high social discount rate downplays the importance of problems like climate change, for which the most significant costs will not arise for decades. Economists Nicholas Stern's and William Nordhaus's widely divergent estimates of the projected costs of climate change are driven by their choices of discount rates.¹¹ Under Stern's

7 ADLER & POSNER, *supra* note 1, at 36–38.

8 MICHAEL LIVERMORE & RICHARD REVEZ, RETAKING RATIONALITY: HOW COST-BENEFIT ANALYSIS CAN BETTER PROTECT THE ENVIRONMENT AND OUR HEALTH 77 (2008).

9 SOC. SEC. ADMIN., ACTUARIAL LIFE TABLE, <https://www.ssa.gov/oact/STATS/table4c6.html> (last visited June 23, 2018).

10 LIVERMORE & REVEZ, *supra* note 8.

11 William D. Nordhaus, *A Review of the "Stern Review on the Economics of Climate Change"*, 45 J. ECON. LITERATURE 686, 689-90 (2007); Cass R. Sunstein & David Weisbach, *Climate*

preferred discount rate of 1.4% per year, the present value of one dollar of damages in one hundred years is twenty-five cents—more than fifty times the present value under Nordhaus's preferred discount rate of 5.5%.¹²

Pro-regulatory advocates like Richard Revesz and Michael Livermore argue that the intergenerational context is “fundamentally different from the context of individual discounting” and an “unavoidably moral decision that should be governed by notions of fairness and equal consideration of all people.”¹³ They support this position, in part, by arguing that the rationale for discounting based on the anticipated greater wealth of future generations is undermined by the negative consequences of climate change projected to be suffered most acutely by individuals in poor, developing countries—like Bangladesh, which has a per capita gross national product (GNP) less than one ninetieth of U.S. GNP and is unlikely to exceed current U.S. income levels within a century.¹⁴ However, because individuals tend to value the welfare of foreigners and abstract future generations less than their own future welfare, this rationale for embracing a low discount rate still requires overriding individual preferences.¹⁵ It is natural for legal and policy advocates to rely on preference-based rationales that support their favored policies and to find reasons why measures of individual preferences should not dictate policy when it would support outcomes they disfavor. Nonetheless, it is worth acknowledging the link between these disputes and how the logic of both the pro- and anti-regulatory positions on valuing life are difficult to reconcile with their positions on time discounting.

Recognizing this tension leaves one with four options. First, one could recommit to a thoroughgoing effort to maximize an aggregation of individual preferences, as inferred by the fallible available tools. This is what Matthew Adler and Eric Posner call “textbook CBA.”¹⁶ Second, one could continue to rely on individual preference-based justifications for policy choices when congenial to one's prior commitments, while explaining away deviations on an ad hoc basis. This is what advocates and scholars pushing both pro- and anti-regulatory agendas tend to do. Third, one could point to the heuristics and biases literature as a justification for eschewing reliance on individual preferences entirely. This approach includes number of alternatives to CBA, including technology-based and other feasibility standards, safety maximization standards, and intuitive balancing all avoid reliance on individual preferences.¹⁷ A special case of this option that retains both a con-

Change and Discounting the Future: A Guide for the Perplexed, 27 YALE L. & POL'Y REV. 433, 433-34 (2009).

12 Mark Harrison, *Valuing the Future: The Social Discount Rate in Cost-Benefit Analysis* 7 (Gov't Productivity Comm'n, Visiting Research Paper 2010).

13 LIVERMORE & REVESZ, *supra* note 8, at 111.

14 *Id.* at 111–12.

15 In fact, it is standard practice outside the domain of climate change for CBA not even to consider the non-domestic costs and benefits of regulation. U.S. OFFICE OF MGMT. & BUDGET, Circular A-4 15 (2003).

16 ADLER & POSNER, *supra* note 1, at 62–63, 187. Adler and Posner also characterize this approach as the “sum-of-CVs test” where CVs refers to contingent valuations, the subjective utility associated with a possible future state of the world.

17 Frank Ackerman & Lisa Heinzerling, *Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection*, 150 U. PA. L. REV. 1553 (2002); ADLER & POSNER, *supra* note 1, at 73–81.

nection to individual subjectivity and the goal of maximizing (a conception of) social welfare is WBA, which relies on present subjective well-being evaluations instead of forward-looking preferences.¹⁸ Finally, one could develop a principled and consistent approach to policy evaluation that neither ignores nor completely defers to measures of individual preferences.

The remainder of this article articulates and defends a vision of the latter option while remaining open to version of WBA. Regulators should not rely on measures of individual preferences in cases when there is no rational basis for that preference that is also normatively viable. I will argue that the only rational basis for pure time preference, Parfitian ideas about personal identity and connectedness over time, severs the connection with social welfare. Since maximizing social welfare is the normative basis for CBA, policymakers should not incorporate preference patterns like pure time preference that do not actually track social welfare.¹⁹ While a broader range of rationales are offered for older people exhibiting similar WTP to avoid near-term mortality risk as younger people, I will argue that this case similarly offers no rationally plausible basis that maintains the link with social welfare. I will also defend CBA against a range of other approaches that eschew reliance on individual preferences and other quantitative proxies for social welfare entirely.

IV. DISCOUNTING THE FUTURE

Discount rates applied to future regulatory costs and benefits should be lowered to eliminate, or at least substantially curtail, the component attributable to pure time preference. Time preference is the degree to which people care less about benefits or costs further in the future.²⁰ Since this preference often partially reflects both the uncertainty associated with future costs and benefits and anticipated decreases in the utility of future consumption, the pure rate of time preference can be defined as the degree to which people care less about a fixed (in probability-adjusted terms) amount of utility in the future relative to current utility.²¹

The general concept of time preference includes two uncontroversial components.²² First, any delayed gratification risks an intervening event that prevents the reward from being reaped. In the individual case, the risk of premature death offers rational support to a perfectly selfish agent to adjust the value of saving for retirement downward in propor-

18 John Bronsteen, Christopher Buccafusco, & Jonathan S. Masur, *Well-Being Analysis vs. Cost-Benefit Analysis*, 62 DUKE L. J. 1603 (2013).

19 See ADLER & POSNER, *supra* note 1, at 25.

20 Shane Frederick, George Loewenstein, & Ted O'Donoghue, *Time Discounting and Time Preference: A Critical Review*, 40 J. OF ECON. LITERATURE 351, 352 (2002).

21 *Id.*

22 *But see* Tyler Cowen & Derek Parfit, *Against the Social Discount Rate*, in JUSTICE BETWEEN AGE GROUPS AND GENERATIONS 144 (James S. Fishkin & Peter Laslett, eds., 1992) (arguing that uncertainty about the future and the declining marginal utility of income should be treated directly, rather than through the social discount rate).

tion to the risk.²³ In the social case, the risk of an asteroid collision causing human extinction reduces the expected benefit of climate change mitigation. Second, the delayed benefit may confer less utility later than now. Inflation will likely make a given dollar value worth less in the future. Even inflation-adjusted dollars may yield less utility if you expect yourself or your society to be richer in the future, due to the declining marginal utility of wealth and income.²⁴

Finally, the capacity of an individual or a society to enjoy a benefit may diminish for reasons unrelated to inflation or economic growth. This is easier to see in an individual case, where a less luxurious vacation taken while one still has the vigor of youth may generate more enjoyment than a more luxurious, but also more sedentary, vacation thirty years later. Accounting for these factors—risk, inflation, and declining marginal utility—is objectively justified.²⁵ The expectation that a specific non-monetary benefit will be less enjoyable in the future is subjective, but seems plausible in some cases. Stripping these components away yields the pure rate of time preference, the rate at which future utility is discounted for no other reason than that it is in the future.²⁶

In the following sections, I consider three arguments about the role of pure time preference in discounting. Section A addresses Derek Parfit's ideas about personal identity and connectedness, which offer a plausible rational justification for pure time preference on the individual level. However, like Parfit, I conclude that this justification is normatively unsustainable on the social level.²⁷ Section B considers opportunity cost as a justification for including pure time preference in the social discount rate and concludes that this rationale only applies to investment-displacing costs. It argues that the Office of Management and Budget's current guidance on discounting cannot be justified even for investment-displacing costs. Section C addresses the argument that pure time preference should be included in the discount rate to maintain coherence with the overall policy regime regarding future orientation. It concludes that this consideration can justify partial incorporation of pure time preference into discount rates applied to consumption-displacing costs.

A. A PARFITIAN DEFENSE

On the individual level, a plausible rationale for pure time preference is that the future person named Gabriel Weil, whose body shares physical continuity with mine and who remembers making (many of) my current choices, is not really or fully me in the

23 Although individual discounting due to risk of non-realization of delayed benefits is uncontroversial and rational, it is not clear that it should be included in the discount rate. If someone saves for retirement, but then dies of a heart attack at age 50, those savings are not lost to society. So, there is a divergence between the individually rational optimal savings rate and corresponding discount rate, and the socially optimal discount rate. Calculations of the costs and benefits of rules also generally account for the expected number of people who will be affected by them, thereby incorporating individual mortality risks. Thus, retaining this component of the discount rate risks double-counting. An analogous concern is discussed in greater detail in Part IV on valuing mortality risks.

24 R. Layard, G. Mayraz, & S. Nickell, *The Marginal Utility of Income*, 92 J. PUB. ECON. 1846 (2008).

25 Frederick, Loewenstein, & O'Donoghue, *supra* note 20.

26 *Id.*

27 DEREK PARFIT, *REASONS AND PERSONS* 170 (1984).

relevant sense. In this view, saving for retirement or forgoing unhealthy foods is more like idiosyncratic generosity toward a specific future person than it is like ordinary self-regarding behavior.²⁸ Parfit wrote, “My concern for my future may correspond to the degree of connectedness between me now and myself in the future . . . since connectedness is nearly always weaker over long periods, I can rationally care less about my further future.”²⁹ Because the vast majority of people discount the welfare of others at least somewhat relative to their own welfare, pure time preference would be a logical extension of that inclination. To the extent that the beneficiary of my financial savings and diet is a different person from my present self, pure time preference is no more open to rational criticism than moderate selfishness. My future “self” may “regret” that my choices favor my current welfare over his, but this would not be so different from a beggar who regrets the insufficient generosity of wealthy people who pass him on the street. While total unselfishness is considered morally praiseworthy, few people would condemn moderate selfishness as irrational. In fact, a common argument against utilitarianism is that it is too demanding in its impartiality, imposing moral duties that individuals cannot be reasonably expected to fulfill.³⁰

The connectedness rationale is not mere philosophical speculation. Studies show that the degree of subjective connectedness with one’s future self-influences savings behavior.

Specifically, when the future self shares similarities with the present self, when it is viewed in vivid and realistic terms, and when it is seen in a positive light, people are more willing to make choices today that may benefit them at some point in the years to come.³¹

28 *Id.*

29 *Id.* at 313. Parfit distinguishes discounting based on connectedness from traditional time-based economic discounting, indicating that his analysis supports the former. He even suggests that discounting would be insignificant in over brief time periods, since people are very closely connected to the yesterday and tomorrow versions of their selves, in contrast to selves separated by forty years. This would be inconsistent with empirically observed discount rates, where are often hyperbolic, showing higher time discount rates in the short term than for more distant delays. However, my intuition is that connectedness aligns well with observed time preference. My self, at this moment, is as connected to me as possible, whereas there is some drop off for my self tomorrow, in part because sleep has severed the continuity of consciousness at least once. By contrast, one can hardly notice a drop off in connectedness when comparing one’s connection to one’s self forty years in the future compared to forty years and one day in the future. Both may feel quite alien, but the latter’s marginal decrease in expected connectedness seems tiny, if even noticeable.

30 See Brian McElwee, *Demandingness Objections in Ethics*, 67 PHIL. Q. 84 (2017).

31 Hal E. Hershfield, *Future self-continuity: how conceptions of the future self transform intertemporal choice*, 1235 ANN. N.Y. ACAD. SCI 30 (2011). See also Jeremy N. Bailenson, Laura L. Carstensen, Jesse Fox, Daniel G. Goldstein, Hal E. Hershfield, William F. Sharpe, & Leo Yeykelis, *Increasing Saving Behavior Through Age-Progressed Renderings of the Future Self*, 48 J. MARK. RES. S23 (2011); Brittany M. Christian et al., *Saving for Your Future Self: The Role of Imaginary Experiences*, 16 SELF & IDEN. 384 (2017).

There is also evidence that people exhibit reduced time preference when make savings decisions for others, rather than themselves.³² This makes sense if pure time preference is driven by attenuated feelings of connectedness with one's future selves.

However, the rationale for individual pure time preference cannot justify relatively high social discount rates. Even if we assume that personal identity in the relevant sense is not stable over time and moderate selfishness is rationally justifiable, this offers no basis for law and public policy to favor the interests of present or near-future person-moments over further future person-moments. Just as we generally expect law and public policy to be impartial when adjudicating the interests of presently existing people, the default presumption should be that law and public policy are neutral with regard to conflicts of interest between present and future versions of Gabriel Weil that will likely exist over the coming decades. Whether or not such neutrality entails paternalistic interventions with respect to "self"-regarding choices on behalf of our future "selves," it certainly seems to imply that public policy should not take present people's valuation of the future at face value for the purpose of evaluating non-paternalistic regulatory interventions. Moreover, there is evidence that individuals exhibit lower discount rates when making decisions on behalf of others.³³ This suggests that people would not want their rate of pure time preference as applied to personal utility to be applied to public policy choices that will affect the welfare of many people. It is also consistent with the interpretation that self-regarding time preference is due, at least in part, to attenuated connectedness with one's future self over time. Thus, the only plausible rational justification for pure time preference severs its connection with social welfare. Since CBA's normative justification is enabling maximization of social welfare,³⁴ preference patterns that fail to track social welfare should be disregarded by policymakers. Likewise, if the Parfitian account of pure time preference fails, this preference pattern is simply irrational and not fit to serve as the basis for public policy.³⁵ This analysis suggests that the component of the discount rate attributable to pure time preference should be subtracted out, increasing the orientation of regulatory policy on the future.

B. OPPORTUNITY COST

What about opportunity cost and market rates of return? OMB Circular A-4, which provides guidelines for CBA, offers three rationales for discounting:

- (a) Resources that are invested will normally earn a positive return, so current consumption is more expensive than future consumption, since you are giving up that expected return on investment when you consume today.
- (b) Postponed benefits also have a cost because people generally prefer present to future consumption. They are said to have positive time preference.

32 Jeremy Shapiro, *Discounting for You, Me and We: Time Preference in Groups and Pairs* (2010), ECON. DEP'T, MASS. INST. OF TECH., <http://economics.mit.edu/files/6059> (last visited July 20, 2018).

33 *Id.*

34 ADLER & POSNER, *supra* note 1.

35 See Shane Frederick, *Discounting, Time Preference, and Identity* (1999) (Ph.D. thesis, Dept. Soc. & Decision Sci., Carnegie Mellon U.).

(c) Also, if consumption continues to increase over time, as it has for most of U.S. history, an increment of consumption will be less valuable in the future than it would be today, because the principle of diminishing marginal utility implies that as total consumption increases, the value of a marginal unit of consumption tends to decline.

There is wide agreement with point (a). Capital investment is productive, but that point is not sufficient by itself to explain positive interest rates and observed saving behavior. To understand these phenomena, points (b) and (c) are also necessary. *If people are really indifferent between consumption now and later, then they should be willing to forgo current consumption in order to consume an equal or slightly greater amount in the future.* That would cause saving rates and investment to rise until interest rates were driven to zero and capital was no longer productive. As long as we observe positive interest rates and saving rates below 100 percent, people must be placing a higher value on current consumption than on future consumption.³⁶

Rationale (c) is discussed above as the declining marginal utility of consumption and should be uncontroversial so long as per capita income can be expected to continue growing. It is worth noting that rising incomes may also be expected to lead people to increase their monetized values of non-market goods like clean air.³⁷ Rationale (b) certainly includes pure time preference, but *may* also incorporate future benefit realization risk and the potential diminution of capacity to enjoy specific benefits discussed above. If the risk of individual non-realization of deferred benefits is a significant component of individual time preference, this may lead to double counting since CBA directly accounts for uncertainties in the magnitude of costs and benefits, including the number of people who will enjoy the benefits. Existential risks that would prevent anyone from experiencing delayed costs or benefits may be best accounted for in the discount rate if realized. In any case, the Circular's language makes clear that rationale (a) is not an independent basis for discounting. Instead, positive risk-adjusted market rates of return are taken as evidence for rationales (b) and (c). If pure time preference is not a proper component of the social discount rate, then by Circular A-4's own reasoning, pure time preference should be subtracted out of any market measure used to determine the discount rate.

But if resources invested at the market rate of return could produce greater benefits in the future, is that not the best outcome under the prevailing constraints? If public policy cannot change individual time-welfare tradeoff decisions to eliminate pure time preference—at least without pervasive and heavy-handed intervention or state control throughout the economy—then shouldn't government only make investments that have higher expected returns than private investments? Perhaps, but not all the costs of regulation come out of investment. Some costs displace present consumption.³⁸ Rejecting the validity of pure time preference would seem to imply different treatment of these costs. Arguably, current practice already makes such a distinction:

36 U.S. OFFICE OF MGMT. & BUDGET, *supra* note 15, at 32 (emphasis added).

37 Nemat Shafik, *Economic Development and Environmental Quality: An Econometric Analysis*, 46 OXFORD ECON. PAPERS 757 (1994).

38 U.S. OFFICE OF MGMT. & BUDGET, *supra* note 15, at 32.

As a default position, OMB Circular A-94 states that a real discount rate of 7 percent should be used as a base-case for regulatory analysis. The 7 percent rate is an estimate of the average before-tax rate of return to private capital in the U.S. economy. It is a broad measure that reflects the returns to real estate and small business capital as well as corporate capital. *It approximates the opportunity cost of capital, and it is the appropriate discount rate whenever the main effect of a regulation is to displace or alter the use of capital in the private sector. . . .*

The effects of regulation do not always fall exclusively or primarily on the allocation of capital. *When regulation primarily and directly affects private consumption (e.g., through higher consumer prices for goods and services), a lower discount rate is appropriate.* The alternative most often used is sometimes called the “social rate of time preference.” This simply means the rate at which “society” discounts future consumption flows to their present value. If we *take the rate that the average saver uses to discount future consumption as our measure of the social rate of time preference*, then the real rate of return on long-term government debt may provide a fair approximation. Over the last thirty years, this rate has averaged around 3 percent in real terms on a pre-tax basis. For example, the yield on 10-year Treasury notes has averaged 8.1 percent since 1973 while the average annual rate of change in the CPI over this period has been 5.0 percent, implying a real 10-year rate of 3.1 percent.³⁹

While OMB Circulars A-4 and A-94 endorse applying different discount rates to consumption and investment, they appear to do so for the wrong reason. Circular A-4 explicitly claims to “take the rate that the average saver uses to discount future consumption as our measure of the social rate of time preference.”⁴⁰ This implies taking observed rates of time preference, based on rationales (b) and (c), including the pure time preference component, at face value as a measure of social time preference.⁴¹ As discussed above, this cannot be justified.

The use of returns on government debt as a proxy for average saver time preference is also unclear, since few savers invest exclusively or even primarily in government debt.⁴² Moreover, most of the federal public debt is held by foreign countries and institutional investors, rather than domestic private savers.⁴³

Perhaps OMB Circular A-4’s approach can be rehabilitated by thinking of interest rates on federal debt in terms of the government’s cost of capital rather than as a reflection of individual time preference.⁴⁴ In this view, government should undertake regulatory interventions and make other investments (e.g., infrastructure) when the expected social returns are greater than the borrowing costs, so long as the up-front costs come out of consumption rather than higher-yield investments. This does not seem to go far enough, for two reasons. First, the government’s borrowing costs are influenced by pure

39 *Id.* at 33 (emphasis added).

40 *Id.*

41 *Supra* text accompanying note 36.

42 Kimberly Amadeo, *Who Owns the U.S. National Debt?*, THE BALANCE, <https://www.thebalance.com/who-owns-the-u-s-national-debt-3306124> (last updated Oct. 28, 2019).

43 *Id.*

44 See U.S. OFFICE OF MGMT. & BUDGET, *supra* note 15, at 33.

time preference, suggesting they offer an overestimate of the true social discount rate. Second, buyers of government debt have the alternatives of both present consumption and riskier investments like real estate, equity markets, corporate debt, and municipal debt.⁴⁵

This suggests that marginal investors are indifferent between the low-risk, low-return value proposition of government debt and higher-risk, higher-return market alternatives. Perhaps society at large has reason to be more risk tolerant than the investment community, but these reasons likely stop short of complete risk neutrality. The same declining marginal utility of income logic that supports discounting also supports a social preference for insuring against downside tail risk.⁴⁶ This suggests that even to the extent that the costs of a regulatory intervention or government investment come out of private investment, 7% likely represents too high an estimate of the appropriate discount rate.

C. POLICY REGIME COHERENCE

A related issue is that many government policies not subjected to CBA influence the savings rate. These include tax-exempt or deferred savings programs like IRAs and 401(k)s, corporate profits taxes, inheritance taxes, investment and R&D tax credits, capital gains and dividends taxes, Social Security, Medicare, the mortgage interest deduction, property taxes, non-taxation of the imputed rental value of land, federal budget deficits, and infrastructure investments.⁴⁷ Some of these programs, like IRAs and 401(k)s, are designed to encourage savings.⁴⁸ Others, like federal budget deficits, capital gains taxes, Medicare, and pay-as-you-go Social Security retirement benefits, have the effect of reducing public and/or private savings.⁴⁹ What would constitute a neutral policy with regard to savings rates is not free from controversy. Many economists argue that taxing any investment returns biases people toward present consumption.⁵⁰ In this view,

45 *Basic Types of Investments – Financial Instruments You Should Know*, MONEY INSTRUCTOR (Oct. 11, 2019), <http://content.moneyinstructor.com/775/financial-instruments-know.html>.

46 Martin L. Weitzman, *Fat Tails and the Social Cost of Carbon*, 104(5) AM. ECON. REV. 544 (2014).

47 Paul A. David & John L. Scadding, *Private Savings: Ultrarationality, Aggregation, and ‘Denison’s Law’*, 82 J. POL. ECON. 225 (1974); Chris Carroll & Lawrence H. Summers, *Why Have Private Saving Rates in the United States and Canada Diverged?* (Nat’l Bureau of Econ. Res., Working Paper No. 2319, 1987); Martin Feldstein, *The Effect of Social Security on Private Savings: The Time Series Evidence* (Nat’l Bureau of Econ. Res., Working Paper No. 314, 1979).

48 Karen Dynan, *Proposal 6: Better Ways to Promote Saving through the Tax System*, in 2–3 THE HAMILTON PROJECT: 15 WAYS TO RETHINK THE FEDERAL BUDGET (Brookings Inst., 2013), https://www.hamiltonproject.org/assets/legacy/files/downloads_and_links/THP_15WaysFedBudget_Prop6.pdf.

49 Laurence J. Kotlikoff, *Health Expenditures and Precautionary Savings* (Nat’l Bureau of Econ. Res., Working Paper No. 2008, 1986); Martin S. Feldstein, *Social Security and Private Savings: International Evidence in an Extended Life-Cycle Model* in MARTIN S. FELDSTEIN M.S. & ROBERT P. INMAN, THE ECON. OF PUB. SERVICES (Robert P. Inman & Martin S. Feldstein eds., 1997).

50 Vernon L. Smith, *Tax Depreciation Policy and Investment Theory*, 4(1) INT’L ECON. REV. 80 (1962); Stephen R. Bond, Michael P. Devereux, & Malcolm J. Gammie, *Tax Reform to Promote Investment*, 12(2) OXFORD R. OF ECON. POL’Y 109 (1996).

IRAs and 401(k)s represent a step toward savings neutrality, rather than affirmative government support for higher savings rates.⁵¹ If overall government policy does not promote higher savings rates and counteract pure time preference, it may be argued that failing to apply market indicators to cost-benefit discounting is incoherent. Likewise, if the overall policy regime does promote savings, this would be reflected in the rates of return used to set discount rates, so no further adjustment would be needed.

A similar issue arises regarding the distribution of costs and benefits across the population. Advocates of replacing CBA with WBA, like John Bronsteen, Christopher Buccafusco, and Jonathan S. Masur, argue that wealth effects distort the WTP measures used in CBA:

It has long been understood that the value an individual places on a risk or a benefit will necessarily be affected by that individual's wealth. A millionaire might think nothing of paying \$10,000 to breathe slightly cleaner air, but someone who must support a family on \$25,000 per year will be much more hesitant to make the same trade-off. Similarly, wealthy people rarely take high-risk jobs because the wage premium is worth less to them and is insufficient to compensate them for the risk. The reason is not that the benefit or risk involved is greater for the wealthier person (though there may be slight differences). Rather, wealth effects are driven by the fact that the money is worth less to the wealthy person. Because cost-benefit analysis involves translating harms and benefits into dollars, these "wealth effects" will affect cost-benefit calculations.

Wealth effects play a large and undeniable role in wage-premium studies, yet CBA cannot fully account for these effects. The fact that rich and poor people (who presumably care equally, or at least comparably, about staying alive) would be willing to pay vastly different amounts to avoid a 1-in-10,000 risk of death illustrates the inadequacy of this metric for valuing lives. WBA circumvents these issues entirely by valuing lives based on individuals' own assessments of their well-being.⁵²

Some ambiguity lies in the claim that rich people and poor people care equally about staying alive. We can grant that this is true in an absolute sense⁵³ while recognizing that if it were possible to transfer the societal resources that would be spent on a stronger rule to the affected population as cash, their WTP, as mediated by wealth effects, is highly relevant. A lot depends on whether you take the sum of societal resources redirected from rich people to poor people as fixed. Advocates for WBA acknowledge this issue, but maintain that agencies should aim to maximize well-being within the scope of their authority:

Any time a government agency must decide between two projects—or two locations for the same project—one of which will affect wealthy people and the

51 See, e.g., Kevin A. Hassett & R. Glenn Hubbard, *Tax Policy and Business Investment*, in HANDBOOK OF PUB. ECON. (Alan J. Auerbach & Martin Feldstein eds., 2002).

52 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1652.

53 As Cass Sunstein points out, even this claim is more questionable in considering "Easy Cases" where the population benefitting from reduced mortality risks from a proposed regulation would also be responsible for paying its cost, as in most regulations of product safety. Sunstein, *supra* note 3, at 40, 61.

other of which will affect poor people, it risks being led astray by wealth effects if it looks at the actual populations of people who will be affected. It may be led to believe that the “wealthy” project will have a greater effect on welfare than the “poor” project, simply because of the impact of wealth on willingness to pay. When the agency cannot tax and transfer—and nearly all agencies lack that authority—it will err and select the wrong project. WBA, on the other hand, would not be confused by wealth effects.

WBA does not require that costs and benefits be translated into dollars, and so the wealth of the affected population cannot confound the analysis. . . .

Economists generally believe that it is more efficient to allocate resources via taxes and transfers than through regulations and new policy proposals. Accordingly, agencies should concentrate on maximizing aggregate wealth and consumption, and welfare and distributional concerns should be left to the tax system. If agencies were to switch to a welfarist decision procedure such as WBA, they would be measuring the wrong quantity.⁵⁴

The authors go on to reject the economists’ critique of WBA, pointing to a number of methodological issues with CBA, which are discussed in Part V.⁵⁵ However, they do not truly grapple with the issue of how an agency should behave when it lacks the tools to achieve the globally optimal outcome.⁵⁶

Three approaches to this issue seem plausible. Approach 1: Agency analysis accepts the rest of government policy as fixed and optimizes its targeted metric (subjective well-being, net benefits, etc.). Approach 2: Agency adopts the policy it believes would be appropriate as part of an optimal overall policy regime. Approach 3: Agency views the level of redistribution and/or future orientation of policy as a matter for democratic deliberation and matches its weighting of these concerns to be in conformity with the overall status quo policy regime, including taxes and transfers. We can treat these three approaches as ideal types, recognizing that the best approach may be a weighted average of two or three.

The first two approaches play out somewhat differently as applied to wealth effects and discounting. For wealth effects, Approach 1 might mean adopting a regulatory posture that weighs the interests of low-income people more heavily than existing tax and transfer policy, meaning a Pareto improvement might be possible by reducing the stringency of a regulation that primarily benefits poor people and compensating them with cash that they value more than the forgone regulatory benefits. Under Approach 2, by contrast, an agency might conclude that most redistribution should be done through taxes and transfers and mostly take WTP measures at face value. There would be some scope for adjustment for wealth effects under this approach. After all, if low-income people had higher after-tax-and-transfer income, they would likely value health and

54 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1654 (emphasis added).

55 *Id.* at 1654–55.

56 They also fail to raise the related concern that projects that concentrate their net benefits on poor neighborhoods may induce changes in property values that price poor people out of those neighborhoods. In principle, however, these effects could be incorporated into either WBA or a form of CBA that incorporates distributional weights. ADLER & POSNER, *supra* note 1, at 143–44.

safety more in monetary terms. It is also dubious that the optimal policy regime would conduct 100% of its redistribution of benefits through taxes and transfers rather than regulatory policy. Nonetheless, these adjustments would likely stop far short of the full cancellation of wealth effects advocated for under WBA.

For discounting, the differences between Approach 1 and Approach 2 would be less sharp. Under Approach 1, regulators would distinguish regulatory costs that come out of present consumption from those that come out of investment, using the status quo cost of capital for investment-displacing costs and a normative, pure time preference-free, discount rate for consumption-displacing costs. Under Approach 2, agency behavior would depend on a judgment regarding the merits of nudges and more coercive forms of paternalism. If the ideal policy regime would include robust employment of nudges (and more coercive interventions when necessary), agencies could simply discount using the market equilibrium cost of capital that they calculate would prevail if people did not exhibit pure time preference, treating present consumption and investment equally. However, if the ideal policy regime includes some special deference to individual autonomy, as maintained by *Nudge* critics, then differential treatment of displaced savings and investment would likely persist as well.⁵⁷

Approach 3 would apply similarly to both cases. With respect to redistribution, there is a tradeoff on some margins between promoting equality and maximizing efficiency.⁵⁸ The prevailing tax and transfer regime and a range of other policy choices reflects a social choice regarding how to balance these competing objectives. This picture is complicated somewhat by the existence of policies that both increase inequality and reduce growth.⁵⁹ Nonetheless, agencies could adopt a weighting of monetized benefits that seeks to align with the implicit democratic choice balancing these competing goals. Similarly, a range of government policies related to public and private savings and investment is reflecting an implicit discount rate. Agencies could simply adopt a discount rate consistent with this overall policy regime and apply it to costs that displace both consumption and investment. Alternatively, if public and private savings and investment are treated sufficiently differently in the broad sweep of public policy, an agency might also infer at least so commitment to the principle of noninterference with private autonomy appealed to by *Nudge* critics.⁶⁰ This might justify continuing to apply a lower discount rate to consumption-displacing costs than to investment-displacing costs, even under an approach that uncritically embraces the prevailing level of future orientation embedded in the policy regime.

How should agencies decide which of these approaches to adopt? First, the viability of Approach 3 is necessarily limited by any specific guidance offered in the statute the agency is administering. Any obligation to act in conformity with democratic choices inferred from a policy regime would be trumped by domain-specific statutory provisions.

57 David F. Bradford, *Constraints on Government Investment Opportunities and the Choice of Discount Rate*, 65 AMER. ECON. REV. 887 (1975).

58 Edgar K. Browning & William R. Johnson, *The Trade-Off Between Equality and Efficiency*, 92(2) J. POL. ECON. 175, 199 (1984).

59 See generally BRINK LINDSEY & STEVEN M. TELES, *THE CAPTURED ECONOMY: HOW THE POWERFUL ENRICH THEMSELVES, SLOW DOWN GROWTH, AND INCREASE INEQUALITY* (2017).

60 Aneil Kovvali, *Who Are You Calling Irrational?*, 110 NW. U. L. REV. 707, 712 (2016).

Nothing in the Constitution requires the Congress to be consistent in its value judgments, and the vagaries of legislative sausage-making routinely reflect temporary and issue-specific coalitions rather than an ideologically coherent program. However, regulatory statutes are typically silent on the issues of discounting and the balance between efficiency and equality.⁶¹ Given the typically general language of regulatory statutes and the Supreme Court's doctrine of *Chevron* deference, agencies have wide discretion in implementing CBA.⁶² It is also unclear that the plethora of economic policies that do address questions of future orientation and of balancing efficiency and equality reflect genuine democratic deliberation over these issues.⁶³ However, an agency may risk political backlash if it pursues an agenda for redistribution of benefits or future orientation that goes beyond what the public is prepared to support. Agencies would be wise to give some weight to the implicit balance of efficiency and equality as well as present and future welfare struck in the legislative process, even if the Congress has not expressed a clear or consistent will on the matter.

Seeking coherence with the overall existing policy regime may be considered a valuable end, independent of concerns about democratic legitimacy. With regard to wealth effects, Approach 3 might be the only way for an agency, given the limited scope of its powers, to avoid contributing to a Pareto inefficient outcome.⁶⁴ This would occur if some population would value a cash transfer more than a regulatory benefit, and the taxes required to fund that transfer are less costly to the taxed parties than the regulation would be. If the agency ignores coherence considerations and proceeds with regulation, the overall outcome would be Pareto inferior to one in which the regulation is repealed and replaced with a tax and transfer adjustment. However, the tax and transfer scheme and other, less transparent, forms of redistribution are sufficiently complex and incoherent that Pareto improvements are likely to be (at least theoretically) possible even if agencies prioritize this sort of policy regime coherence over other considerations.⁶⁵

It is also not clear that when looking at two alternatives that are Pareto non-comparable (e.g., regulate or not), the theoretical existence of a Pareto improvement to one of those options is relevant.⁶⁶ These considerations apply with greater force to the discounting domain where it is not well established that paternalistic interventions to promote future orientation are preferable to adjustments of non-paternalistic regulation toward greater temporal neutrality. Thus, while policy coherence may be a legitimate

61 See Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 446 (1989).

62 See *Chevron U.S.A., Inc. v. Nat. Res. Def. Council, Inc.*, 467 U. S. 837 (1984).

63 See William N. Eskridge, Jr., *Politics without Romance: Implications of Public Choice Theory for Statutory Interpretation*, 74 VIRG. L. REV. 275, 276–77 (1988).

64 A Pareto inefficient outcome is one for which it is possible to make at least one person better off without making anyone worse off (a Pareto improvement). HAL R. VARIAN, *INTERMEDIATE MICROECONOMICS: A MODERN APPROACH* 15 (8th ed. 2010).

65 Deborah L. Paul, *The Sources of Tax Complexity: How Much Simplicity Can Fundamental Tax Reform Achieve*, 76 N.C. L. REV. 151, 167–68 (1997).

66 Pareto non-comparability means that moving from one alternative to the other makes at least one person better off and at least once person worse off. David W. Pearce, *THE DICTIONARY OF MODERN ECON.*, 331 (2d. ed. 1983).

objective to which some weight should be given, both in its own terms and on democratic legitimacy grounds, we can reject a pure implementation of Approach 3.

As indicated above, the choice between the Approach 1 and Approach 2 is much more significant for wealth effects than for discounting. Neither Approach 1 nor Approach 2 would apply the pure time preference component of discounting to consumption-displacing costs. The major difference is regarding the treatment of investment-displacing costs and only arises if you accept the legitimacy of some amount of paternalism regarding savings rates and other choices involving time tradeoffs. Having rejected pure implementation of Approach 3, we can therefore conclude that the existence of sub-optimally utilized policy levers affecting savings rates cannot justify full inclusion of pure time preference in discount rates for consumption-displacing costs. Pure time preference may only be included in the discount rate for consumption-displacing costs to the extent that we give some weight to Approach 3 and conclude that the overall existing policy regime favors present consumption over future benefits to a degree that is consistent with affirming pure time preference.

V. VALUING LIFE

Let us now return to the question of how to value mortality risks. The currently dominant practice in the United States is population-average VSL, which places a fixed value on a statistical life regardless of age or health status.⁶⁷ The VS LY approach would instead value mortality risks based on an estimate of how many life-years would be lost from the expected deaths, weighing preventing the death of a young and healthy person as a greater benefit than preventing the death of an older or sicker person.⁶⁸ Population-average VSL also differs from textbook VSL, which allows for valuations that are heterogeneous by income, age, and other attributes.⁶⁹ It is worth quoting Revesz and Livermore's defense of population-average VSL and critique of the VS LY alternative at some length.

The life-years method does not flow from either sound economic theory or good facts. *The approach is fundamentally inconsistent with the important tenet of economic theory in which value is determined by the willingness to pay.* Under that tenet, the economic value of mortality risk reductions should be determined by how much an individual would voluntarily exchange for the reduction. It would only be economically defensible to decrease the value assigned to mortality risk reduction to account for age if one's willingness to pay decreases as one ages.

But the life-years method ignores willingness to pay as a proxy for value, and instead assumes a downward linear relationship between a person's age and the value of that person's life. *This assumption is inconsistent with the standard economic observation that individuals generally assign greater value to goods that are more limited*

67 Cass R. Sunstein, *Lives, Life-Years, and Willingness to Pay*, 104 COLUM. L. REV. 205, 205–206 (2004). Different agencies do use different values for statistical lives, but for each agency the VSL method uses a fixed value regardless of age.

68 LIVERMORE & REVESZ, *supra* note 8.

69 ADLER & POSNER, *supra* note 1.

in supply. The technique uses a constant per life-year value, so that all life years are valued equally no matter when they occur during the life cycle.

As people age, they can anticipate fewer future life years. Because of this scarcity, we might expect that they value their future life years more highly than younger people would. *By assuming that no difference exists between the values a 40-year-old and a 70-year-old would attribute to an additional year of life, the life-years method overlooks the effect of scarcity on valuation.* By ignoring the effect of scarcity and focusing regulatory efforts on reducing risk for young and healthy people, the life-years method delivers regulatory benefit to those who value it least. This approach takes the standard economic logic of “willingness to pay” and stands it on its head. Generally, the most efficient system is the one that moves resources to the people that value them most. The life-years method accomplishes exactly the opposite. Moreover, across a certain age range of their lives, *as people grow older, they have more income and wealth. It is well established that the willingness to pay to avoid risk is highly correlated with income. The greater affluence of middle-aged individuals (at least preretirement) thus suggests an increase in willingness to pay,* contrary to the prediction of the life-years method.

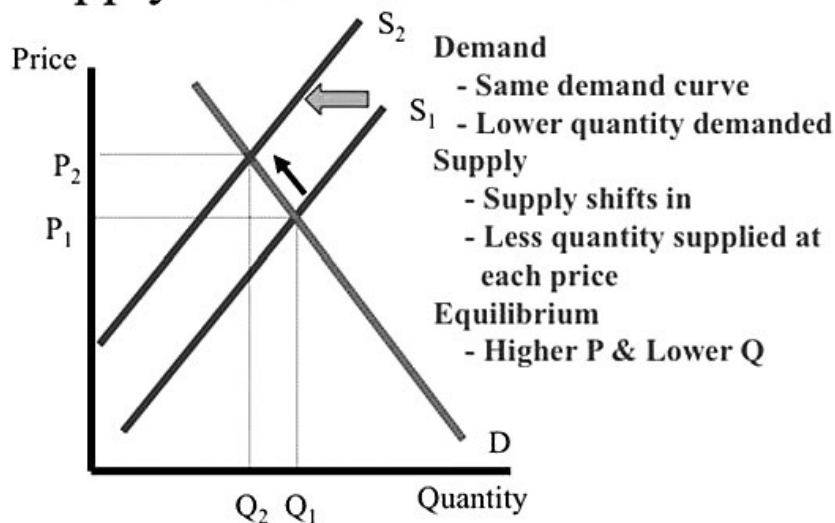
Various models—all ignored by life-years advocates—seek to determine how the value of risk reduction might change with age. *Some models predict that as the probability of death increases, so does the willingness to pay to avoid risk, because people cannot take money to their graves.* In other models, increases in background risk, which occur as people age, decrease the willingness to pay for a specific risk. Other models are simply ambiguous. It is possible that none of these models captures the whole story. What is important, however, is that no plausible economic model offers even lukewarm support for the diminishing linear relationship between life expectancy and willingness to pay that undergirds the life-years method. The life-years method, then, is entirely without theoretical justification.⁷⁰

Let us assume, for the sake of argument, that WTP to avoid near-term mortality risks is constant throughout the life cycle. That is, a thirty-year-old American woman with a remaining life expectancy of fifty-two years exhibits no greater WTP to avoid a near-term mortality risk than a sixty-three-year-old woman with a remaining life expectancy of twenty-two years.⁷¹ Under this assumption, population-average VSL converges with textbook VSL, at least with regard to age. This is the most favorable assumption we can make for population-average VSL method—that individual preferences offer no basis for deviating from a constant VSL over the life cycle. If the case for the VSL method falls short under this assumption, we can safely reject it.

70 LIVERMORE & REVESZ, *supra* note 8, at 80–81 (emphasis added).

71 SOC. SEC. ADMIN., *supra* note 9.

Supply Shifts Left



Revesz and Livermore offer two principal reasons why roughly constant WTP to avoid mortality risk across the life cycle might be plausible: scarcity and wealth effects. Regarding scarcity, they claim that “individuals generally assign greater value to goods that are more limited in supply.”⁷² This claim is somewhat misleading. In orthodox economic theory, the shape of the demand curve is independent of supply.⁷³

What is true is that a leftward shift in the supply curve will tend to raise both the price and the marginal value of a unit of a good, even as the total⁷⁴ quantity demanded falls.⁷⁵ This is a consequence of the tendency of demand curves to slope downward—for consumers to buy less of a good as the price rises.⁷⁶ This, in turn, is a function of diminishing marginal utility.⁷⁷ A tenth apple is worth less to a hungry person than the first, so she will only buy ten if the price is sufficiently low. If she is hungry, however, her willingness to pay for the first apple may be quite high. In this sense, the first apple can be said to be more valuable to her than the tenth. If apples become scarcer, the price will rise and she will buy fewer, but the marginal value of the last apple she buys will be greater than when apples are plentiful. If this translates at all to the context of life-years, it suggests that the marginal value of a life-year *decreases* over the life cycle. That, at least, is what the concept of diminishing marginal utility would imply. It is true, however, that not all demand curves slope down monotonically.⁷⁸ In this context, there may

72 LIVERMORE & REVESZ, *supra* note 8, at 80.

73 N. GREGORY MANKIW, *PRINCIPLES OF ECONOMICS* 82 (2009).

74 BRIGHAM YOUNG UNIV.-IDAHO, *ECON 150 Economic Principles and Problems – Micro: Section 01: Supply and Demand*, https://courses.byui.edu/econ_150/econ_150_old_site/lesson_03.htm (last visited Aug. 12, 2018).

75 *Id.*

76 *Id.*

77 George J. Stigler, *The Adoption of the Marginal Utility Theory*, 4 *HIST. OF POL. ECON.* 571, 579 (1972).

78 See, e.g., Harvey Leibenstein, *Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand*, 64(2) *Q. J. OF ECON.* 183 (1950).

be reasons to think that the first year of life is not the most valuable. Indeed, charity evaluator GiveWell weighs lost life-years against factors like level of cognitive function at the age of death.⁷⁹ Nonetheless, the true claim that scarce goods tend to be more valuable to those who consume them simply does not support the conclusion that years that fall later in the life cycle are more valuable than earlier years.

With regard to wealth effects, note that Revesz and Livermore are essentially claiming that policymakers should value benefits to (pre-retirement) older people more highly precisely because older people have more income and wealth.⁸⁰ This is in tension with their argument, discussed in Part II, that we should not apply standard economic discounting to climate change because many of those most affected by its impacts will be poor Bangladeshis. I sincerely doubt that Revesz and Livermore would favor broad application of the idea that regulatory efforts should be focused on reducing risk for wealthy and high-income people, but this is precisely what their reasoning implies.⁸¹ Moreover, even if CBA should value the lives of wealthier people more because of their higher WTP, then it would be better to do so directly rather than using age as a proxy for wealth and income, as Revesz and Livermore's rationale for retaining population-average VSL implies.⁸² Using a fixed VSL not only imperfectly tracks actual wealth and income levels, but also entangles wealth effects with irrational or normatively unsustainable drivers of WTP like pure time preference.

Revesz and Livermore's argument that older people's WTP are driven up by the fact that they "cannot take money to their graves," should be particularly unpersuasive to regulators.⁸³ To the extent this is the correct explanation for older people's increasing WTP per life-year, what is changing is not the value of life-years as people age; it is the value of money. Money may not have much value to a person who is about to die, but the money (and real resources) invested in lowering mortality risk will necessarily be diverted from the living population. Individuals may not be able to take money to the grave with them, but that does not mean that money and the real resources it commands lose their social value when the person dies. Thus, to the extent that older people's WTP to avoid mortality risks is driven up by this phenomenon, the link between their preferences and social welfare is broken. As with pure time preference, this break negates the normative value of the preference pattern and renders it irrelevant to policymakers.

Let us consider the issue from a moral perspective.⁸⁴ From an outside perspective, the death of a thirty-year-old woman who would have lived another fifty-two years seems to be worse than the death of a sixty-three-year old woman who otherwise would have lived another twenty-two years.⁸⁵ There is room for reasonable disagreement with regard to

79 Robert Wiblin, *Finding the Best Charity Requires Estimating the Unknowable*, 80,000 HOURS (July 16, 2018), <https://80000hours.org/podcast/episodes/james-snowden-givewell-research/>.

80 LIVERMORE & REVESZ, *supra* note 8.

81 Cass Sunstein points out that the case for valuing mortality risks for wealthier and higher-income populations more highly is strongest in the "Easy Cases" when the beneficiaries of a proposed regulation would also be responsible for paying its costs. See Sunstein, *supra* note 3, at 40, 61.

82 LIVERMORE & REVESZ, *supra* note 8, at 81.

83 *Id.*

84 *Id.* at 110–11.

85 For further discussion on this point, see Sunstein, *supra* note 67, at 214–16.

whether the average value of a life-year in one's thirties, forties, and fifties is higher or lower than in one's sixties, seventies, and eighties. As one ages and loses the vigor and health of youth, she is compensated with wisdom, financial security, and grandchildren.⁸⁶

A full appraisal might consider average happiness and life satisfaction over the life-cycle, as well as the contributions individuals make to society. Studies have generally shown a U-shaped happiness curve, with people reporting high life satisfaction in early adulthood that gradually declines before recovering as they enter retirement.⁸⁷ However, recent work has called into question whether this result is an artifact of individuals' reporting functions.⁸⁸ A reasonable case can be made that earlier or later life-years are marginally more valuable. What cannot be credibly argued is that the three years between ages eighty-two and eighty-five that a sixty-three-year-old woman would lose to premature death are approximately equivalent in value to the thirty-three years between ages thirty and sixty-three that a thirty-year-old woman would lose. (They both would lose the nineteen years between ages sixty-three and eighty-two, in expectation).⁸⁹ However, this is precisely the view one must adopt in order to conclude that the age of the potential beneficiary of mortality risk reduction is irrelevant, as the VSL method does.

Revez and Livermore fail to mention another potential explanation for older people valuing near-term mortality risk reduction at similar levels to younger people: time preference. In particular, the Parfitian idea that pure time preference is rationally justified by decreasing levels of connectedness with successive future selves seems highly relevant.⁹⁰ In this view, when younger people consider near-term mortality risks, they are giving little weight to potential lost life-years decades in the future, when the person who would or would not exist is not fully "them." As with other manifestations of Parfitian pure time preference, the rational defensibility of this behavior does not imply that policymakers should uncritically accept this revealed preference as the basis for non-paternalistic regulation.⁹¹ As argued above, public policy should be neutral between the

86 See generally ATUL GUWANDE, *BEING MORTAL: ILLNESS, MEDICINE AND WHAT MATTERS IN THE END* (2014).

87 Timothy N. Bond & Kevin Lang, *The Sad Truth About Happiness Scales: Empirical Results* (Nat'l Bureau of Econ. Res., Working Paper No. 24853, 2018).

88 *Id.* at 19–20.

89 Technically, life expectancy tables compress a wide distribution into an average. Someone who is currently thirty years old is no more likely to die at eighty-two than someone who is currently sixty-three. The reason she has a lower expected death age is that she might die before age sixty-three. This does mean she is somewhat less likely than a sixty-three-year-old to experience the life-years between eighty-two and eighty-five, but much less than a naive interpretation of life expectancy tables would imply. However, this consideration only amplifies the absurdity of treating the sixty-three-year-old's remaining twenty-two expected life-years as equally valuable to the thirty-year-old's fifty-two expected remaining life-years, since the overlap in life stage of the lost life-years is even greater than the nineteen years that the naive calculation suggests.

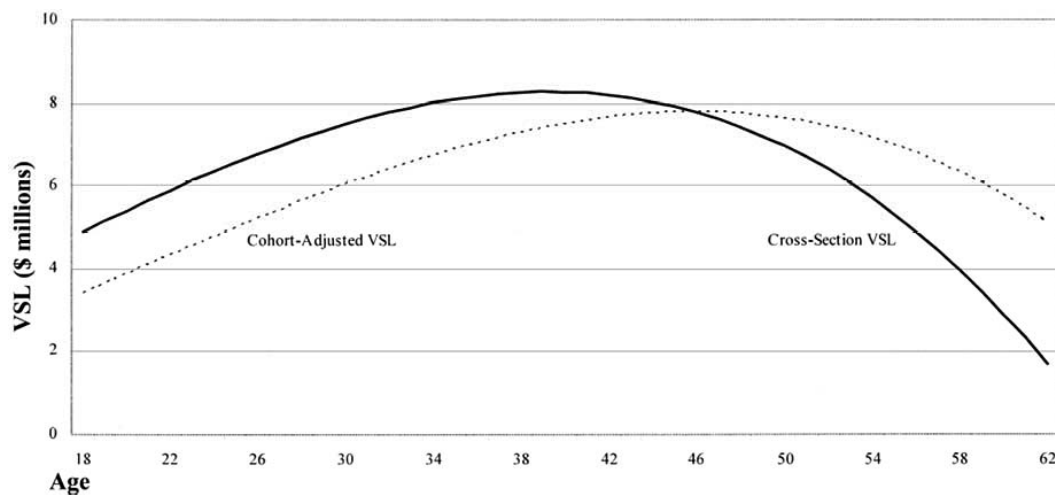
90 PARFIT, *supra* note 27.

91 Non-paternalistic regulation refers to any regulation designed primarily to regulate other-regarding activities like pollution or product safety, as opposed to self-regarding activities like diet, exercise, and retirement savings. Most major rules subject to CBA are non-paternalistic in this sense.

interests of present and future iterations of a citizen, just as it should be between different contemporary citizens. Thus, as with pure time preference, no rationale for the pattern of individual preferences that might support the population average VSL approach is rationally plausible while also maintaining the connection to social welfare that normatively justifies deference to individual preferences.

Joseph Aldy's and W. Kip Viscusi's analysis of Bureau of Labor Statistics Census of Fatal Occupational Injuries data yield an inverted U-shaped pattern for VSL over the lifecycle.⁹²

FIGURE 1—COHORT-ADJUSTED AND CROSS-SECTION VALUE OF STATISTICAL LIFE, 1993–2000



Viscusi notes that annual earning and consumption patterns “exhibit a trajectory that that mimics the overall shape of the VSL-age relationship.”⁹³ He then claims that “This similarity is not a statistical quirk, as there is a theoretical linkage of one’s valuation of safety to income and consumption levels.”⁹⁴ Viscusi stipulates that, in an idealized economic model where people were endowed with their lifetime wealth at birth, VSL would steadily decline with age, corresponding to a roughly fixed VSLEY.⁹⁵ Given imperfect credit and insurance markets, however, individuals’ financial resources change significantly over their lifecycle.⁹⁶ Given this empirical relationship, Viscusi rejects proposals to scale back valuations of mortality risk based on the remaining life expectancy at risk, because “such quantity adjustments are not based on how people’s willingness to pay for the risk reduction varies across these policies.”⁹⁷

92 Joseph E. Aldy & W. Kip Viscusi, *Adjusting the Value of a Statistical Life for Age and Cohort Effects*, 90(3) REV. ECON. & STAT. 573 (2008).

93 W. KIP VISCUSI, *PRICING LIVES: GUIDEPOSTS FOR A SAFER SOCIETY* 99–100 (2018).

94 *Id.*

95 *Id.* at 97–98.

96 *Id.*

97 *Id.* at 104.

In addressing the relative fairness of population average VSL and VSLY, Viscusi considers an extreme case comparing a thirty-year-old with a remaining life expectancy of forty-nine years to an elderly person with advanced emphysema and a life expectancy of one year.⁹⁸ He concedes that using a population average VSL for valuing risks to both of their lives is inequitable in ignoring the extra forty-eight years of life expectancy at risk for the thirty-year-old.⁹⁹ However, he claims that the VSLY approach “also creates a bias by steadily reducing the total benefit attached to the remaining years of life as a person ages.”¹⁰⁰ It is not clear why this represents a bias. All people will die at some point and lose their remaining life expectancy. All public policy can ever do is influence the likelihood that they will die at any time. Viscusi compares both the population average VSL and VSLY unfavorably to textbook VSL, but holds that “pinpoint matching of the person’s valuation of the risk reduction to particular policies is generally not possible.”¹⁰¹ He concludes population average VSL is efficient and equitable for policies with broad impact and that VSLY should be reserved for “special instances in which the policy delivers only a minor effect on life expectancy” and thus “utilization of the VSLY approach retains a linkage to private valuations of risk.”¹⁰²

Viscusi’s approach is based on the fundamental premise that willingness to pay is the ultimate arbiter of value. It is indeed rational for willingness to pay for a good to rise with wealth and income, as the marginal value of a dollar falls. Given political or economic (including negative incentive effects) constraints on redistribution, there may be a case for adjusting the value of mortality risks for wealth and income. If “safety” were the *same good* across the lifecycle, it would make sense to defer to individual preferences about it that track willingness to pay for other consumption items. However, reductions in mortality risk buy a person fewer expected life-years as they get older. To the extent that individual willingness to pay purely tracks wealth and income across the lifecycle, it ignores the quantity of life lost, suggesting a myopic focus on the near future. Thus, while it may make sense to adjust VSLY for wealth and income, the variation in available financial resources over the lifecycle is not a good reason to ignore the diminishing quantity of left at risk as a person ages. This point is underscored by Aldy and Viscusi’s method of extracting VSLY estimates across the lifecycle, shown in figure 2 below.¹⁰³

98 *Id.*

99 *Id.*

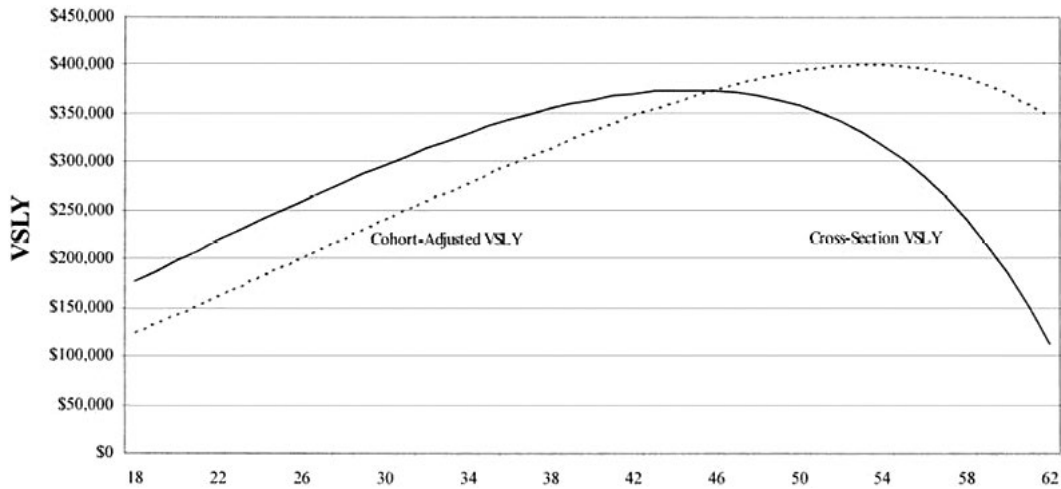
100 *Id.* at 108.

101 *Id.*

102 *Id.* at 109.

103 Aldy & Viscusi, *supra* note 92, at 579.

FIGURE 2—VALUE OF A STATISTICAL LIFE-YEAR BASED ON COHORT-ADJUSTED AND CROSS-SECTION VALUE OF STATISTICAL LIFE, 1993–2000



Aldy and Viscusi apply a 3% discount to the remaining expected life years to back out an estimate of VSLY by age from VSL figures.¹⁰⁴ Recall that life expectancy already incorporates other mortality risks that might detract from the value of reducing any specific risk, so the risk that one might not realize the full benefit of later saved life-years cannot justify applying this discount rate.

The expectation that society will be wealthier in the future cuts both ways in this context, since wealthier future people can be expected to value their lives more than present people by Viscusi's own reasoning. Indeed, Viscusi implicitly acknowledges as much in his discussion of discounting future mortality risks, where he suggests applying a 3% discount rate, but offsetting it with a 2% annual increase in VSL, for a net discount factor of 1%.¹⁰⁵ If anything, the case for discounting is stronger in the case of future mortality (as opposed to later life-years lost due to current mortality risks), because benefit realization risk may not be separately accounted for in that case. Applying a 3% discount rate to statistical life-years thus leans very heavily on pure time preference, which lacks a normatively compelling rationale. If the safety the variation in VSL over the lifecycle were driven entirely by wealth and income effects, as Viscusi suggests, a 3% discount rate would not be needed to produce an age-VSLY curve that roughly track to willingness to pay for other goods.¹⁰⁶

Life valuation does raise an additional complication. Traditional advocates of the VSLY method favor taking the VSL calculated using a population of average age forty and dividing that figure by the life expectancy of a forty-year-old to derive the value of a life-year.¹⁰⁷ If the forty-year-olds take insufficient account of the value of life-years that successive versions of their selves will experience in the coming decade, this method will

104 *Id.* at 579–80.

105 VISCUSI, *supra* note 93, at 128–29.

106 *Id.*

107 LIVERMORE & REVESZ, *supra* note 8, at 77.

underestimate the value of a life-year. In this sense, Viscusi, Revesz, and Livermore are correct that existing VS LY methods are naïve and inadequate. On the other hand, older people focused on the idea that they “cannot take money to their graves” may overestimate the value of a life-year, at least in monetized terms. A rigorous application of the VS LY approach to valuing risk reduction will require further methodological refinement to identify an appropriate life-year value. These methodological challenges, however, do not justify continued application of the population average VSL method, which relies on preferences that, if rationally justifiable at all, are not normatively valid inputs for CBA. The upshot of shifting from the population-average VSL method to a modified VS LY approach would be to weigh the deaths of middle-age and younger people more heavily and the deaths of older people less heavily in evaluating proposed regulations.

VI. WELL-BEING ANALYSIS

Perhaps the issues raised about relying on individual preferences in policy analysis justify a sharper break with the current practice of CBA. One alternative is WBA, which seeks to maximize individual self-reported life satisfaction, measured in well-being units (WBUs), instead of the satisfaction of preferences in monetized units as CBA does.¹⁰⁸ The case for making this switch hinges on whether how well a person’s life is going is best judged by how often they get what they want or how happy they are. I address this debate in Part VI. But first, Part V argues that although WBA advocates like Bronsteen, Buccafusco, and Masur offer a trenchant critique of CBA, they rightly acknowledge that WBA is not intended to solve every problem. Of interest here, difficult decisions will still have to be made about valuing life and time discounting, thorny issues that will need to be addressed whether or not the switch is made from CBA to WBA.

A. MORTALITY RISK IN WBA

Advocates of WBA point out several sources of error associated with the revealed preference studies frequently relied upon in CBA. This critique is at its strongest in discussing informational and computational problems. Bronsteen, Buccafusco, and Masur note that wage premium studies assume, without adequate justification, that workers are well informed about the risks of various occupations, are able to assimilate low mortality risks (e.g., 1 in 10,000) so as to affect job choice and wage bargaining, and act on information about mortality.¹⁰⁹ They also argue that a 1-in-10,000 mortality risk may be “too fine-grained for regression analysis to detect” in the context of a choice between jobs that vary on a number of dimensions other than wages and safety.¹¹⁰ Given these challenges, they argue that it should not come as a surprise that wage premium studies have produced VSL estimates ranging from \$100,000 to \$76,000,000.¹¹¹ W. Kip Viscusi points out that observed variation in VSL plausibly tracks genuine differences between populations and contexts, rather than measurement issues.¹¹² For instance, VSL esti-

108 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1618.

109 *Id.*

110 *Id.* at 1648.

111 *Id.* at 1650.

112 VISCUSI, *supra* note 93, at 104.

mates under \$1 million come from studies of very high risk occupations with annual fatality rates of 1/1000.¹¹³ It stands to reason that people with comparatively low valuations of mortality risks will self-select into these jobs.¹¹⁴ The authors discuss wealth effects and affective forecasting errors, which cast doubt on the normative value of preferences based on expectation of the impact of outcomes on later well-being.¹¹⁵ These are indeed serious methodological challenges for CBA, and Bronsteen, Buccafusco, and Masur claim that WBA “sidesteps nearly all of these problems.”¹¹⁶

WBA advocates also argue that traditional CBA is wrong to treat all deaths as equally bad, at least in welfarist terms. Accordingly, they agree that the VSLY method is an improvement over the VSL approach.¹¹⁷ They correctly point out that both approaches ignore the quality of the life-years saved.¹¹⁸ The quality-adjusted life-years (QALY) approach seeks to address this shortfall, but WBA proponents argue that the methods used to elicit QALY values overstate the negative utility associated with many negative health states, among other issues.¹¹⁹ In particular, they suggest that time trade-off and standard gamble studies using healthy people will tend to underweight lost life-years compared with years of impaired health.¹²⁰ However, WBA also faces significant challenges with the relative weighting of mortality and other risks. Consider Bronsteen, Buccafusco, and Masur’s application of WBA:

b. Cancer Cases Avoided. The EPA provided a range of estimates for the number of cases of cancer that will be avoided under each regulatory option. In the interest of simplicity, we base our calculations on the median number. There are limited available data on the welfare loss that an individual experiences when she is sick with cancer, but one study calculated the welfare loss from “stomach/liver/kidneys or digestive problems,” which we believe is the closest analog. *That welfare loss is 0.238 WBUs per year while the person is sick. We assume that the typical individual who dies from cancer caused by dioxin and furan effluents is sick with cancer for two years and then dies thirty years before she normally would. This is obviously a rough assumption, but it is no rougher than the EPA’s assumption that all lives are equivalently valuable and have a median value of \$5.75 million. The average American has a life satisfaction of 7.4 (again, on a scale of 0.0 to 10.0). When an individual dies, she loses all of the welfare that she might otherwise have experienced throughout the remaining years of her life.* Thus, we calculate the welfare benefit from avoiding one fatal case of cancer by the following equation:

(2) Welfare benefit from avoided fatal cancer = 2 x (0.238 WBUs) + 30 x (7.4 WBUs) = 222.48 WBUs.¹²¹

113 *Id.*

114 *Id.* at 8.

115 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1618.

116 *Id.* at 1650–51.

117 BRONSTEEN, BUCCAFUSCO, & MASUR, HAPPINESS & THE LAW 85 (2015) [hereinafter HAPPINESS & THE LAW].

118 *Id.*

119 *Id.* at 86–87.

120 *Id.* at 87.

121 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1641 (emphasis added).

Whatever methodological and conceptual flaws CBA and QALYs may be said to suffer from, this application of WBA cannot be considered an improvement. In this example, the authors treat not being alive as equivalent to being alive with a life satisfaction of 0.0, the lower bound of the scale used.¹²² That equates not being alive to the worst imaginable torture.¹²³ None of us know for sure what happens when we die, but it is reasonable to assume that death represents the end of subjective experience. That is, it is like nothing to be dead, much like it is like nothing to not have been born. It is a neutral experience (or non-experience, if you prefer). Most people's lives may be better than neutral,¹²⁴ but it is an enormous leap to conclude that *any* time spent alive is necessarily above the neutral point.

To be fair, Bronsteen, Buccafusco, and Masur do envision WBA using a negative ten to ten scale with zero as the neutral point.¹²⁵ Accordingly, treating years not spent alive as worse than any possible living years is not inherent in WBA. They use a zero to ten

122 One potential concern about this mode of analysis is its implications for the welfare impact of population growth. A straightforward application of the conception of WBA presented here would imply that adding a new member to the population is to be favored up until the point that his existence reduces the well-being of others by more than his total expected well-being. This reasoning, of course, leads to Parfit's repugnant conclusion, which holds that no matter how large a population of ecstatic people you have, there will always be a much larger population of people whose lives are just barely worth living that should be preferred because it has greater aggregate welfare. PARFIT, *supra* note 27, at 381–90. There are three potential responses to this. The simplest is to embrace the repugnant conclusion, arguing that our intuitive aversion to it reflects scale insensitivity. Second, one could argue that WBA only applies to presently existing people. This seems arbitrary, however, since WBA would judge any future state by the well-being of then-existing people, regardless of whether they existed at the time at which any prospective analysis was done. Finally, WBA advocates could argue that the same conundrum applies to CBA. This may be correct. However, CBA seems to be on somewhat stronger ground in claiming that not-presently-existing people cannot be said to have preferences. Therefore, new births would only affect CBA to the extent that existing people have preferences about them. However, CBA does purport to be able to analyze the costs and benefits of policies whose primary effects are for future generations, which suggests it is using the preferences of current people as proxies for the welfare of future people. CBA defenders could still claim that these preference proxies do not include the desire to be born, since this is not a preference that anyone can have. In any case, the repugnant conclusion is a long-standing and vexing puzzle in population ethics, and it may be unreasonable to expect WBA to solve it. This footnote simply serves to flag the issue for future discussion.

123 Bronsteen and Masur both dispute this characterization, saying that they explicitly point out that WBA should be conducted using a negative ten to ten scale with a zero-neutral point. They say it is unfair to judge the merits of WBA based on this first example, which must rely upon imperfect existing data, particularly when comparing it to current CBA practice, which is the product of decades of refinement. Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1633; E-mail correspondence with John Bronsteen, Jonathan S. Masur, and Gabriel Weil, Fellow, Univ. of Cal. Irvine School of Law, Ctr. for Land, Env't, and Nat. Res. (Aug. 14, 2018) (on file with author).

124 *But see* DAVID BENATAR, BETTER NEVER TO HAVE BEEN: THE HARM OF COMING INTO EXISTENCE (2008).

125 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1618.

scale in the sample analysis because most existing data was collected using that scale.¹²⁶ In that case, zero is not an appropriate neutral point. We do not know for sure what point subjects assume corresponds to neutrality on a zero to ten scale, but zero is not a plausible assumption. Their analysis also implies that a year of not being alive is more than thirty-one times worse than a year of being sick with cancer, meaning that a person would be better off living with cancer for thirty-two years before finally succumbing than she would be living thirty-one healthy years and then dying quickly and painlessly. This seems implausible.

Faced with this challenge, WBA advocates emphasize that they are not seeking to maximize preference fulfillment, as CBA does, but to maximize well-being. People may predict that they will lose a great deal of welfare when suffering from cancer, but studies of people suffering from similar symptoms do not always bear this out.¹²⁷ People underestimate their capacity to adapt to new conditions and thus might overestimate the welfare loss from non-fatal cancer.¹²⁸ However, even if we grant that the goal of public policy should be to maximize happiness rather than preference fulfillment, there is reason to doubt that self-reports of life satisfaction, particularly for cancer patients, are reflective of their actual welfare. They may be affected by social desirability bias and/or try to look brave and convince themselves and their loved ones that things are not so bad.¹²⁹ People may also have systematically different reporting functions.¹³⁰ In particular, what is interpreted as hedonic adaptation may instead be an adaptation of individual utility scales such that the threshold for reporting the highest levels of life satisfaction is lowered.¹³¹ Unlike the random variations in reporting functions that Bronsteen, Buccafusco, and Masur argue will tend to wash out in large populations averages,¹³² this sort of systematic shift in response to health or wealth/income shocks would present more significant methodological hurdles for WBA.¹³³ This may be just as significant a source of error as the affective forecasting errors that WBA advocates point to for wage-premium studies. As Bronsteen, Buccafusco, and Masur rightly point out, however, these limitations must

126 *Id.* Bronsteen, Buccafusco, and Masur's argument that it is not fair to judge the potential of WBA based on a sample analysis conducted with flawed data is also in some tension with their claim that "[t]here is already a treasure-trove of longitudinal data on life satisfaction that has been collected over the decades in the United States, Great Britain, and Germany." *HAPPINESS & THE LAW*, *supra* note 117, at 51.

127 Shane Frederick & George Loewenstein, *Hedonic Adaptation*, *WELL-BEING: FOUNDATIONS OF HEDONIC PSYCHOLOGY* 302–29 (1999).

128 *Id.*

129 See Robert J. Fisher & James E. Katz, *Social Desirability Bias and the Validity of Self-Reported Values*, 17 *PSYCH. & MARKETING* 105, 116 (2000). Social desirability bias refers to the tendency of survey respondents to answer questions in a manner that will be viewed favorably by others. It poses a serious problem with conducting research using self-reports.

130 Bond & Lang, *supra* note 87, at 43–46.

131 *Id.* at 45–46.

132 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1626–27.

133 Bond & Lang, *supra* note 87, at 43–46; Richard E. Lucas, *Adaptation and the Set-Point Model of Subjective Well-Being Does Happiness Change After Major Life Events?*, 16 *CURRENT DIRECTIONS IN PSYCH. SCI.* 75 (2007).

be assessed in comparison to those of CBA.¹³⁴ Resolving this methodological debate is beyond the scope of this article.

WBA advocates might also accept some of my other criticisms of this specific implementation of WBA and try to adjust the method accordingly. For instance, they might accept the need for some tweaks to their method for assessing the well-being loss from cancer, but maintain that this information is much more reliable than assessments based on people's predictions of how bad cancer will be, as inferred from studies of wage differentials confounded by a multitude of uncontrolled variables. Indeed, Bronsteen has privately expressed openness to tweaks along these lines, while maintaining that self-reported life satisfaction is a better proxy for welfare than WTP.¹³⁵ As for the utility loss from death, Bronsteen, Buccafusco, and Masur do recognize the need to conduct new studies with a negative to positive range where zero is explicitly identified as representing an equal balance between positive and negative experience, equivalent to not being alive.¹³⁶

It is likely that a workable form of WBA could be devised along these lines. It may even be an improvement over CBA. It is not plausible, however, to claim that WBA neatly sidesteps the methodological issues associated with mortality risk in CBA. Perhaps the most significant obstacle facing WBA is that it is not possible to ask non-living people how happy they are. This is especially important because prolonging life dominates non-fatal health problems, monetary costs, and other considerations under the vision of WBA presented by Bronsteen, Buccafusco, and Masur. Getting this feature wrong could easily wipe out any advantages associated with other aspects of WBA. To weigh the benefit of prolonging life against other policy goals, a methodology must compare the value of being alive to being dead; but an experiential measure like WBA makes that impossible because no one can report on the experience of death. Even using a negative ten to ten scale with death as the zero point, there would still be no way to know how death compares to life in terms of experiential reporting—that is, in terms of the method of valuation that WBA uses. Is death half-way between the worst experience and the best one? How could one possibly know, or even guess, given WBA's methodology? One option would be to instruct participants in happiness studies to treat the zero rating as equivalent to not being alive. Of course, this would require participants to *imagine* what it is like not to be alive and expose the method to the same concerns about affecting forecasting errors that WBA advocates raise about CBA.

Bronsteen acknowledges that the impossibility of eliciting the life satisfaction of non-living people presents a challenge for WBA, but does not think it offers a reason to prefer any other methodology.¹³⁷ On my reading, this is dependent on his claim that “alternatives like CBA do such a bad job of measuring welfare by *any* plausible yardstick of what welfare is.”¹³⁸ If CBA could overcome the methodological challenges associated with measuring individual preferences about mortality risk *and* those preferences are a

134 HAPPINESS & THE LAW, *supra* note 117, at 92.

135 E-mail correspondence with John Bronsteen and Gabriel Weil, Fellow, Univ. of Cal. Irvine School of Law, Ctr. for Land, Env't, and Nat. Res. (Aug. 14, 2018) (on file with author).

136 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1618.

137 E-mail correspondence with John Bronsteen and Gabriel Weil, Fellow, Univ. of Cal. Irvine School of Law, Ctr. for Land, Env't, and Nat. Res. (Aug. 14, 2018) (on file with author).

138 *Id.*

plausible proxy for welfare, then CBA would have a stronger claim to measuring a quantity of relevance to valuing life. Perhaps CBA falls sufficiently short on one or both of those criteria to outweigh this limitation of WBA. Nonetheless, WBA's commitment to present subjective experience evaluation as the best proxy for welfare does render the impossibility of collecting these data for the single most important effect of regulation (especially in WBA's own weighting) a particularly pressing challenge for WBA.¹³⁹ Bronsteen agrees that more work is needed in this area and suggests this challenge may be addressed within WBA by looking at the happiness ratings reported by people when they are numb as a proxy for not being alive.¹⁴⁰ He also suggests the WBA might borrow some preference-based methods from CBA, saying:

[A]lthough the best way (for reasons we provide at length) to gauge the welfare-value of an experience is to record people's in-the-moment self-reports of that experience, that's not the *only* way. Unlike CBA or any other methodology, WBA uses the best way for everything that it can, and for something like death where that best way is impossible, it uses second-best alternatives. . . . WBA faces this problem with assessing the effect of death on well-being, but other methodologies face this problem with assessing the effect of *everything* on well-being.¹⁴¹

This approach would expose WBA to the same criticisms Bronsteen, Buccafusco, and Masur level at CBA in the context of valuing mortality risks. It would also revive the debate over competing VSL, VSLY, and QALY methods. Combining preference-based measures with subjective well-being measures would also introduce of additional complexity into WBA. Nonetheless, these costs may be worth paying if WBA proponents are correct that self-reports of experience are the best way to measure welfare.

B. TIME DISCOUNTING IN WBA

Regarding time discounting, advocates of WBA claim that since WBA is not based on monetary values, "time value of money" rationales involving interest rates and inflation do not apply to WBA.¹⁴²

Unlike money, well-being is time invariant. Five WBUs in 2021 are worth just as much in welfare terms as 5 WBUs in 2011. Indeed, the entire reason that the value of money varies over time is that the amount of well-being it can be used to purchase varies over time. Thus, there is no need to discount in order to accommodate the time-value of well-being. Many of the difficulties with discounting that force EPA to report results at two different discount rates, and the interagency climate change working group to do so at three different rates, are simply irrelevant to WBA.¹⁴³

In Bronsteen, Buccafusco, and Masur's account, the importance of discounting in CBA is driven primarily using dollars as the unit of account. In this view, accounting for the

139 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1683.

140 E-mail correspondence with John Bronsteen and Gabriel Weil, Fellow, Univ. of Cal. Irvine School of Law, Ctr. for Land, Env't, and Nat. Res. (Aug. 14, 2018) (on file with author).

141 *Id.*

142 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1686.

143 *Id.*

costs and benefits of regulation in terms of WBUs allows policy analysts to mostly avoid the discounting puzzle. I disagree.

To the extent that the compliance costs of regulation displace investment, that will tend to decrease the well-being of the would-be beneficiaries of those investments. Even if one is convinced that increases in wealth as such have minimal effects on well-being, some of that displaced investment is likely to be in technologies that would have extended lives, to which WBA would give great weight. In principle, WBA could try to directly track what the marginal funds would be invested in and trace their impact on future well-being. Analogous direct accounting for displaced investments could be done in CBA, which is precisely what Cowen and Parfit advocate in their case for a zero social discount rate.¹⁴⁴ In practice, however, it is much simpler to calculate the economic returns on investment as a proxy for well-being and then translate those returns into WBA.

All the issues discussed above regarding potential differential treatment of investment-displacing vs. consumption-displacing costs would thus apply to WBA. As discussed above regarding wealth effects, the thorny issue of how to address coherence with government policy not subject to WBA would also remain. Embracing WBA might entail a more paternalistic approach that would imply government policy should promote higher savings rates, but agencies applying WBA will still lack the policy tools to implement that vision.¹⁴⁵ Likewise, exogenous existential risks like asteroid impacts that would wipe out all deferred benefits (whether accounted for in dollars or WBUs) are another driver of discount rates that WBA must account for. Working with a non-monetized measure of value might avoid the need to address adjustments associated with inflation and the declining marginal utility of wealth and income, but even these are likely to creep back when considering investment returns. In any case, discount rates are typically applied to real dollars, with inflation accounted for separately.¹⁴⁶ Thus, WBA cannot avoid most of the messy issues associated with selecting a discount rate, and switching back and forth between dollars and WBUs as the unit of account may introduce some additional complexity.

Bronsteen, Buccafusco, and Masur acknowledge that WBA may have to entertain the possibility of discounting based on individuals' "pure time preferences for immediate gratification over later benefits" and other vaguely specified rationales.¹⁴⁷ As explained

144 See Cowen & Parfit, *supra* note 22.

145 The paternalism claim is not meant as a criticism of WBA. As my analysis of pure time preference suggests, it is plausible that public policy should indeed be more paternalistic in terms of promoting higher savings rates. It is worth noting, however, that if Bronsteen, Buccafusco, and Masur are correct that marginal increases in life expectancy are much more valuable than increases in non-health-related consumption, this suggests a much more comprehensive suite of paternalistic interventions may be warranted. As with future-orientation, one might conclude nudges or more coercive interventions are inappropriate even if individual choices are not entitled to deference in evaluating non-paternalistic regulations.

146 John Whitehead, *Discount Rates for Benefit-Cost Analysis*, CROMULENT ECON. BLOG (August 10, 2005), https://www.envecon.net/2005/08/discount_rates_.html#targetText=Most%20discount%20rates%20used%20for,are%20virtually%20default%20risk%20free.&targetText=Assuming%20expected%20inflation%20is%20equal,%20free%20interest%20is%201.74%25.

147 Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1688.

in OMB Circular A-4, however, the time value of money rationale for discounting overlaps with other plausible rationales, rather than providing independent justification.¹⁴⁸ Part of the explanation for the fact that one dollar a year from now is worth less than a dollar today is indeed the empirical reality of time preference, including its uncontroversial components like future benefit realization risk and the residual pure time preference. It would be an odd result to retain the least defensible component of the discount rate while setting aside less controversial components associated with returns on investments and deferred benefit realization risk.

None of this means that WBA is inherently incompatible with a rigorous and normatively justified approach to discounting. However, it does suggest that WBA must grapple with many of the same questions about the future orientation of public policy that CBA must and possibly some new questions that are unique to WBA. While it may be true that the value of WBUs is time-invariant, this conclusion does not resolve most of the important questions about discounting, including those addressed in earlier parts of this article.¹⁴⁹ It also suggests, as Bronsteen, Buccafusco, and Masur implicitly concede, that more work will be needed before WBA is ready for implementation.¹⁵⁰

Nonetheless, if WBA advocates are correct in arguing that subjective well-being is a better proxy for welfare than preference fulfillment, it is worth refining the methodology and considering its use in policy evaluation.¹⁵¹ It is relevant in this context that a lot of the supposed methodological advantages of WBA relative to CBA, particularly sidestepping affective forecasting errors, come from evaluating CBA by the standard of subjec-

148 U.S. OFFICE OF MGMT. & BUDGET, *supra* note 15, at 32.

149 Bronsteen, Buccafusco, and Masur acknowledge that a role for pure time preference may be retained in WBA. Bronsteen, Buccafusco, & Masur, *supra* note 18, at 1688. Likewise, if future people are expected to have higher well-being (not just wealth), perhaps diminishing marginal returns apply to the non-monetary inputs to well-being. Egalitarian considerations might also favor marginal increments of WBUs for relatively unhappy currently living people over the same increase in WBUs for happier future people. Also, the claim should at least be clarified to refer to certain WBUs. Any realization risk for future WBUs would justify weighting them less heavily than more certain near-term WBUs, whether this risk is accounted for directly or through some sort of discount factor.

150 *Id.* at 1633. This point was also suggested in e-mail correspondence with John Bronsteen. E-mail correspondence with John Bronsteen and Gabriel Weil, Fellow, Univ. of Cal. Irvine School of Law, Ctr. for Land, Env't, and Nat. Res. (Aug. 14, 2018) (on file with author).

151 WBA also differs from traditional CBA in its method of aggregation of individual welfare to social assessment. WBA combines a utilitarian social welfare function method of aggregation with a happiness-based conception of well-being. CBA's method of aggregation, using dollars as the common unit, could also be applied to a happiness-based conception of well-being. See Matthew D. Adler, *Happiness Surveys and Public Policy: What's the Use?*, 62 DUKE L. J. 1509 (2013). Conversely, a utilitarian or prioritarian social welfare function method of aggregation could be applied to a preference-based conception of well-being. Matthew D. Adler, *A Better Calculus for Regulators: From Cost-Benefit Analysis to the Social Welfare Function* 17 (Duke Law School Public Law & Legal Theory Series, Working Paper No. 2017-19, 2017) [hereinafter *A Better Calculus*]. While Adler makes a compelling case that a utilitarian social welfare function would be an improvement over CBA as a method of aggregation, that issue is beyond the scope of this article. For present purposes, the relevance of WBA is as a potential alternative approach to the issues raised above regarding using measures of individual preferences in policy analysis.

tive well-being maximization. It should not be surprising if a policy analysis tool explicitly designed around self-reports of well-being does better by that standard than one designed around preference fulfillment. What people want is not always what is most conducive to their overall life satisfaction.¹⁵² In this respect, WBA is a more paternalistic method than CBA, which generally takes people's preferences at face value.¹⁵³

Lisa Robinson's analysis of WBA applies the opposite frame as Bronsteen, Buccafusco, and Masur's:

WBA does not tell us how individuals prefer to allocate resources. Whereas statistical analysis can be used to estimate the relationship between measures of subjective well-being and income, *such analysis does not indicate whether the affected individuals would willingly exchange income for that level of well-being*. Nor does WBA ask individuals how they would prefer to allocate money across different goods and services, including nonmarket outcomes such as improved health. Rather, *it assumes that individuals would prefer to see resources allocated so as to achieve a higher level of subjective wellbeing, however defined*. Thus, if used as a decision criterion, WBA is more paternalistic: the analyst decides that money should be allocated so as to maximize well-being, even if those affected would prefer to allocate resources differently.¹⁵⁴

Characterizing WBA as if people *prefer* a resource allocation that maximizes well-being applies the standard embodied in CBA to WBA. Each approach requires strong methodological assumptions in order to function as a good approximation of the other. This is true even though both ultimately seek to measure and maximize a conception of social welfare – to make people better off. Applying the standard associated with one's favored approach and then characterizing the alternative method as facing greater methodological challenges in meeting it unnecessarily confuses this issue. Both WBA and CBA face significant methodological challenges, even in terms of the conception of welfare that each is crafted to maximize (preference fulfillment or subjective well-being, within legislative and moral side constraints). When the other method's standard is applied, the methodological and conceptual challenges multiply.

Bronsteen, Buccafusco, and Masur offer compelling arguments for conceiving of social welfare in terms of happiness rather than preference satisfaction or an objective list of goods. The next part of this article considers those arguments alongside Matthew Adler and Eric Posner's restricted preferences account.

152 Daniel J. Benjamin et al., *Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference*, 104(9) AM. ECON. REV. 2698 (2014).

153 Again, this is not necessarily a criticism of WBA. Complete deference to individual preferences is not normatively justified, as earlier parts of this article argue. The question is how much normative value preferences should have—if any—over and above their ability to predict subjective well-being.

154 Lisa A. Robinson, *Cost-Benefit Analysis and Well-Being Analysis?*, 62 DUKE L. J. 1717, 1731 (2013) (emphasis added).

VII. HAPPINESS VS. PREFERENCE SATISFACTION

Adler and Posner offer a book-length defense of CBA.¹⁵⁵ Part of this defense is their contention that maximizing fulfillment of individual preferences, subject to certain conditions, is what it means to maximize welfare.¹⁵⁶ They in turn endorse weak welfarism, under which maximizing social welfare is one of multiple objectives that have moral relevance.¹⁵⁷ Bronsteen, Buccafusco, and Masur similarly allow that non-welfarist considerations like moral rights and fairness may have value independent of their contribution to aggregate welfare.¹⁵⁸ In justifying their approach, Adler and Posner consider two alternative conceptions of welfare: mental-state accounts and objective-good accounts.¹⁵⁹ They claim to demonstrate these views are both inferior to fulfillment of lauded preferences.¹⁶⁰ Consider their analysis of mental-state accounts:

Freud, wracked by pain at the end of his life, refused painkillers because they would have impeded his thinking. He preferred the more painful mental state (thinking clearly but suffering great pain) to the more pleasant mental state (thinking fuzzily but suffering no pain) and, intuitively, was better off with the more painful mental state.

[M]ental state accounts generally, remain vulnerable to a powerful objection. All such accounts insist that welfare is solely a function of our mental states. If *P*'s mental states are identical in O_1 and O_2 , then, regardless of the other ways in which the two outcomes might differ, *P*'s welfare must be the same in both states. This is implausible. Imagine that *P* wants to be reputed to be a great scholar, and has been systematically tricked by his colleagues into thinking that he has that reputation; in fact, they are unimpressed by his scholarship and belittle it behind his back. O_1 is the actual outcome, namely, one in which *P* incorrectly believes he has a good scholarly reputation; O_2 is a counterfactual outcome, in which *P* believes he has a good scholarly reputation and really does (his colleagues admire his work). Then *P*'s mental states are identical in O_1 and O_2 , but, intuitively, O_2 is better for *P*'s welfare than O_1 . No mental-state account can validate this kind of intuition.¹⁶¹

One question we can ask about this is how well the mental-state account maps onto WBA. In Freud's case, the matter would turn on whether Freud's subjective reports of life satisfaction in the two cases tracked his preference to forgo pain killers.¹⁶² If not, WBA would favor ignoring Freud's preference unless doing so violated rights or similar possible side-constraints prohibiting direct autonomy violations. In the case of a scholar

155 ADLER & POSNER, *supra* note 1. Adler has since disavowed CBA as a method of aggregating individual welfare, but remains committed to a preference-based conception of welfare. A *Better Calculus*, *supra* note 151.

156 A *Better Calculus*, *supra* note 151, at 36.

157 *Id.* at 54.

158 HAPPINESS & THE LAW, *supra* note 117, at 162–63.

159 ADLER & POSNER, *supra* note 1.

160 *Id.*

161 ADLER & POSNER, *supra* note 1, at 30.

162 HAPPINESS & THE LAW, *supra* note 117, at 179–80.

with a potentially false impression of having a great reputation, it seems clear that WBA will be indifferent to the reality of the scholar's reputation, so long as he was unable to detect his colleagues' trickery. Indeed, this case is like the deceived spouse case examined by Bronsteen, Buccafusco, and Masur:

Jack is very happy in his marriage to Jill, and fully believes she is faithful to him. In one possible state of affairs, she is cheating on Jack without his knowledge. We are asked to suppose that these two states of affairs are identical but for the cheating, as far as Jack's experience of life is concerned. To wit: Jill treats Jack *identically* whether she is cheating on Jack or not; Jack never learns that Jill is cheating on him; and Jack's experience of life is never affected by the cheating in any way.¹⁶³

Bronsteen, Buccafusco, and Masur go on to consider variations on this hypothetical where Jill's cheating takes place on a business trip to Nepal, closely before or after Jack's death. These variations tend to pump one's intuition closer to the view that happiness, in the sense of positive mental states, offers a complete account of welfare. Bronsteen, Buccafusco, and Masur then offer the following analysis:

We think that people's intuitions about the original example may be driven by their failure to honor the example's rules. This would make sense because it is almost impossible to believe that Jill would treat Jack *identically* if she were cheating and if she were faithful. When we try to imagine those two states in the original example, we picture Jill very differently in each of them. In the state where she is faithful, we picture her loving Jack and having no interest in cheating on him. This picture seems like a recipe for a solid, lasting marriage and for great happiness along the way. But when we picture the unfaithful state, we struggle to imagine Jill acting *exactly* the same way toward Jack. How could she possibly do that unless she were some sort of sociopath, or a pathological liar, or at a minimum a cold and unfeeling person? We think that for most people, their feelings would show through in one way or another. Either they would feel guilty and let it show, or their cheating would be motivated by dissatisfaction that would show, or there would be some other manifestation of their cheating. If Jill really showed nothing, then it means that Jack is married to someone who seems a lot different from our picture of Jill in the faithful state, and that affects our intuitions about how much happiness Jack is getting out of the marriage (notwithstanding the stipulation to the contrary).¹⁶⁴

I suspect that WBA proponents are correct that most people's intuitions are driven by smuggled-in happiness-affecting differences that violate the stipulation of the hypothetical. However, imagine we strip down the differences in the original hypothetical, stipulating that in both cases Jill truly does love Jack, but does not believe that either marital infidelity or deception are immoral (perhaps because she accepts a subjective well-being account of welfare). If she encounters an opportunity to cheat that is attractive, convenient, and low risk, she will take it. The only difference between the case in which she cheats and the case in which she doesn't cheat is whether such an opportunity

163 *Id.* at 167.

164 *Id.* at 168–69.

arises. In this hypothetical, it really does seem plausible that Jack's happiness could be entirely unaffected by this difference, even if their marriage lasts for decades after the cheating occurs. I confess that I lack strong intuitions about whether Jack would be made worse off by such cheating. However, I do think that enough people do have a strong intuition that *objective reality* matters to welfare, over and above its impact on subjective experience, that this stripped-down hypothetical still presents something of an anomaly for the mental state account of welfare.

This *objective reality* intuition is also pumped by Robert Nozick's famous experience machine hypothetical, which Bronsteen, Buccafusco, and Masur also address.¹⁶⁵ Indeed, they argue that in the most important cases, the experience machine essentially reduces to the deceived spouse case.¹⁶⁶ I concur, with the proviso that the experience machine offers the prospect of *perfect* deception, which really could have no impact on happiness, much like my stripped-down version of the deceived spouse hypothetical. In both cases, the residual difference between happiness-equivalent cases relates to the consistency of objective reality with subjective experience.

Another example might be a physicist, say Isaac Newton, whose theories of physics seemed completely accurate during his lifetime, but were later found to be flawed. One could reasonably ask whether he would have been better off had his theories been correct. We can suppose he had a preference to discover true theories of physics, not merely true-seeming approximations. A mental state account of happiness would hold that Newton would have been no better off had his theories been true, so long as this had no observable consequences communicated to him during his lifetime. Perhaps this is correct, but the possibility that objective reality is relevant to welfare, over and above subjective experience, cannot be easily dismissed.

WBA advocates also argue that, even if the hedonic theory leads to a counterintuitive result in the experience machine case, this would not be sufficient to refute an account that covers countless other cases better than any other theory.¹⁶⁷ This argument was more compelling back in the 1970s when Nozick thought up the experience machine hypothetical. Today, the looming prospects of engrossing virtual reality worlds and direct brain stimulation have the potential to create positive subjective mental experiences largely untethered from objective physical reality.¹⁶⁸ As public policy confronts these technologies, it will be important to have an account of welfare that handles deviations between subjective experience and objective reality in a satisfactory way.¹⁶⁹ Perhaps that will mean overriding the intuitions and corresponding preferences of those who believe that positive mental experiences are less valuable when disconnected from true beliefs about external reality. Perhaps not. In any case, it will be increasingly diffi-

165 ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 42–45 (1974). The experience machine would stimulate an individual's brain to provide what pleasurable experience he desires. While plugged into the machine, people would be unable to distinguish their experience from a similar one in the outside world.

166 HAPPINESS & THE LAW, *supra* note 117, at 168–69.

167 *Id.* at 175.

168 See Yew-Kwang Ng, *Towards Welfare Biology: Evolutionary Economics of Animal Consciousness and Suffering*, 10 BIOL. & PHIL. 255, 258 (1995).

169 HAPPINESS & THE LAW, *supra* note 117, at 175.

cult to set aside such concerns as irrelevant and implausible science-fiction hypotheticals.¹⁷⁰

Now consider one final example that Bronsteen, Buccafusco, and Masur address as raising potential objections to their hedonic theory of welfare:

In one state, Jane lives in poverty in a society that discriminates against her based on her sex. But she has adapted to the poverty and discrimination and feels quite happy despite it. In the other state, Jane lives a life of comfort, activity, and commitment to justice, all in a society that gives her full political and social rights and nourishes her capabilities. But she feels no happier throughout her life than if she had lived in poverty in oppression.¹⁷¹

Bronsteen, Buccafusco, and Masur again argue, correctly in my view, that “the example gains all its intuitive traction from telling people to disregard something that they cannot in fact disregard.”¹⁷² While it may be true that we cannot disregard our assumption that someone who is free and rich and able-bodied is happier than someone who is oppressed, poor, or disabled, this is precisely what WBA asks us to do when people report the same life satisfaction levels. Indeed, Bronsteen, Buccafusco, and Masur cite WBA’s superior accounting for hedonic adaptation as a positive virtue relative to CBA. Thus, while they are right that this example does not falsify the hedonic theory of well-being, it does underscore the methodological challenges facing WBA. WBA advocates may respond that CBA must confront adaptive preferences, just as WBA must confront hedonic adaptation. However, at least Adler and Posner, as discussed below, recognize that adaptive preferences do not correspond to welfare and advocate excluding them from CBA (while recognizing this may be infeasible in practice).¹⁷³ Thus, while adaptation poses a challenge for preference and mental state accounts of welfare, it is worth noting that Bronsteen, Buccafusco, and Masur’s formulation of WBA is based on the view that all reported adaptation correspond to happiness and thus welfare. This affords greater relevance to common intuitions that some people are worse off than others who report the same subjective life satisfaction ratings.

There is at least one potential objection to their hedonic theory of welfare that Bronsteen, Buccafusco, and Masur do not fully address: the intrapersonal and intertemporal aggregation problems.¹⁷⁴ Bronsteen, Buccafusco, and Masur’s account of WBA holds that every moment of experience and every increment of life satisfaction should be weighted equally. It also, as discussed above, holds that extending one’s life further in time is equivalent to packing the net positive utility experienced in the extra years into a shorter life.¹⁷⁵ However, their defense of a hedonic account of welfare never actually argues for these propositions. This is despite Bronsteen, Buccafusco, and Masur’s statement that a theory of well-being must answer questions like, “[i]s Sam better off if he

170 *Id.*

171 *Id.* at 170.

172 *Id.* at 171.

173 ADLER & POSNER, *supra* note 1, at 128–29.

174 HAPPINESS & THE LAW, *supra* note 117, at 158–60. Bronsteen, Buccafusco, and Masur do address a special case of intrapersonal aggregation in their discussion of the well-being of a marathon runner.

175 *Id.* at 166–67.

extends his life an extra year by making healthy choices that sacrifice some of his enjoyment of life?"¹⁷⁶ This is indeed an important question, but asking living people about their life satisfaction cannot alone answer it. Living people do, however, have preferences about the tradeoffs between extending their lives and increasing the average utility of the time they are alive. Thus, a preference-based account of welfare offers a subjective anchor for valuing life extension that is unavailable to mental state accounts.

A related question relates to the distribution of positive and negative experiences within a person's life. While this issue relates to time preference, it is analytically distinct. Even from an intemporal perspective, one may prefer a variety of positive and negative moments in her life to a steady stream of mediocre ones, even in a life with the same total net positive affect. Likewise, some people may value increments of positive net affect differently at different points along the well-being scale. For instance, moving from five to six may be more/less valuable than from seven to eight. Preference theories can, however imperfectly, account for these subtleties. It may be that the approach imbedded in Bronsteen, Buccafusco, and Masur's formulation of WBA, treating all increments of net positive affect as equally valuable, including when they are spread over different lengths of time, is better than any alternative available. Nonetheless, their defense of its theoretical underpinnings is incomplete.

An additional complication is that the concept of happiness contains significant internal ambiguities. WBA is based on ratings of overall life satisfaction, but this is just one of many measures of well-being. Other potential measurements include moment-to-moment reports of subjective utility (addressing Kahneman's experiencing self, as opposed to the remembering self),¹⁷⁷ the frequency and intensity of positive emotions, the prevalence of depression, and how meaningful people rate their lives as being.¹⁷⁸

Cross-national studies of these measures show different countries coming out on top depending on the measure used. For overall life satisfaction, Scandinavian countries tend to score highest.¹⁷⁹ Latin American countries rate highest on measures of positive emotion.¹⁸⁰ When it comes to minimizing the prevalence of depression, Australia scores best.¹⁸¹ African countries, by contrast, tend to rate highest when people are asked about

176 *Id.* at 136.

177 Daniel Kahneman & Jason Rills, *Living, and Thinking About It: Two Perspectives on Life*, in *THE SCIENCE OF WELL-BEING* 285 (Nick Baylis, Felicia Huppert, & Barry Keverne eds., 2005).

178 Scott Alexander, *The Tails Coming Apart as Metaphor for Life*, *SLATE STAR CODEX* (Sept. 25, 2018), <http://slatestarcodex.com/2018/09/25/the-tails-coming-apart-as-metaphor-for-life/>.

179 See John F. Helliwell, Richard Layard & Jeffrey D. Sachs, *World Happiness Report 2018* (Mar. 14, 2018) <http://worldhappiness.report/ed/2018/>; Frank Martela, *Finland Is the Happiest Country in the World, and Finns Aren't Happy about It*, *SCI. AM.* (May 11, 2018), <https://blogs.scientificamerican.com/observations/finland-is-the-happiest-country-in-the-world-and-finns-arent-happy-about-it/>.

180 Jon Clifton, *Latin Americans Lead World in Emotions*, *GALLUP NEWS* (Aug. 27, 2015), <https://news.gallup.com/poll/184631/latin-americans-lead-world-emotions.aspx>.

181 Fiona J. Charlson et al., *Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010*, 10(11) *PLOS Medicine* e1001547 (2013).

how meaningful they their lives are.¹⁸² People can reasonably disagree about which of these measures, or what welfare function aggregating across multiple measures, public policy should seek to maximize. Likewise, individuals may have different preferences between moment-to-moment utility, overall life satisfaction, positive emotion, avoiding depression, and leading a meaningful life. Indeed, in addition to happiness and life satisfaction, surveyed people report caring about “other items, such as goals and achievements, freedoms, engagement, morality, self-expression relationships, and the well-being of others.”¹⁸³ Measures of object-level individual preferences, at least theoretically, reflect individuals’ priorities among these meta-level objectives. Collapsing well-being into a single measure substitutes the policymaker’s judgment regarding the proper meta-level objective for the individuals’ judgment. This heightens the burden on advocates of replacing preference-based measures with any particular form of reported well-being to prove that doing so will actually make people better off in the ways that matter most to them.

Now consider Adler and Posner’s critique of objective-good accounts:

Objective-good views of welfare are vulnerable to the following criticism: O_1 cannot be better for P ’s welfare than O_2 , if P does not (at some time) prefer O_1 to O_2 . Listening to opera might be, objectively, a better use of someone’s time than watching sitcoms, but unless she prefers opera to sitcoms (at least ex post, having been exposed to opera, if not ex ante) the world in which she listens to opera is not better for her than the world in which she watches sitcoms. Similar examples might be constructed for any objective good. An “objective” good (as we use that term) is necessarily some feature of the world that can occur without the subject preferring it. . . . Thus, *all these accounts overlook the crucial point that each individual is a (partial) sovereign with respect to his own welfare. Something that P doesn’t want for himself, and never comes to want, can’t make him better off.*¹⁸⁴

The last sentence above is the crux of Adler and Posner’s position. It reduces to a bare intuition that preferences trump subjective experience. Their caveat regarding the timing of preferences offers a slight concession to the affective forecasting errors objection of Bronsteen, Buccafusco, and Masur.¹⁸⁵ However, a paraplegic who reports life satisfaction that is not significantly lower than what she reported prior to paralysis may nonetheless maintain a strong preference not to be paralyzed. Thus, while Adler and Posner indicate some flexibility regarding which preferences are relevant to welfare, they are committed to the notion that a person “cannot be made better off in the teeth of her actual preferences.”¹⁸⁶

Given the evidence of a gap between *liking* and *wanting*, however, their position is vulnerable to the objection that the form of CBA they defend will fail to give people

182 Alexander, *supra* note 178.

183 Daniel J. Benjamin et al., *supra* note 152, at 2700.

184 ADLER & POSNER, *supra* note 1, at 32 (emphasis added).

185 See Bronsteen, Buccafusco, & Masur, *supra* note 18.

186 ADLER & POSNER, *supra* note 1, at 36.

what they like, even as it seeks to give them what they want.¹⁸⁷ Indeed, Yew-Kwang Ng argues that the same logic that justifies evaluating outcomes based on informed rather than actual preferences, followed to its logical conclusion, implies that happiness is what matters.

Just as actual preferences should be discounted due to the effects of ignorance and spurious preferences, informed preferences should also be discounted due to some inborn or acquired tendencies to be irrational, such as placing insufficient weights on the welfare of the future, maximizing our biological fitness instead of our welfare.¹⁸⁸

This may not be a fatal objection, but then neither is Adler and Posner's core objection to objective-good and mental-state accounts: they sometimes fail to maximize fulfillment of self-interested preferences that survive idealizations.

Ultimately, therefore, Adler and Posner only succeed in demonstrating that the mental-state, objective-good, and preferentialist accounts are truly three distinct conceptions of welfare that appeal to different intuitions. They offer examples where most people's intuitions suggest that preferences matter over and above happiness, but different cases evoke different intuitions. Ng offers the following counterpoint:

To see that happiness is more fundamental than preference, consider advanced computers in the 21st or 22nd century that have preferences but no affective subjective feelings. Clearly their preferences should not count morally. If it is replied that only human (informed) preferences should count, not machine preferences, then consider animals now and advanced computers in the 25th century that do have subjective affective feelings, i.e. they have pain, joy, etc., then most morally sensitive persons will agree that their welfare should also count. Thus, clearly welfare is more important and fundamental than preferences, informed or not, ultimately speaking.¹⁸⁹

Bronsteen, Buccafusco, and Masur add two important critiques of restricted preference accounts of welfare. First, they note that there exist clear cases, such as self-sacrifice out of a sense of obligation, where people prefer outcomes that decrease their well-being.¹⁹⁰ Moreover, no one has developed an adequate account of how to restrict preferences to self-interested ones that offers clearer guidance than sticking to "preferences about well-being."¹⁹¹ Adler and Posner do argue against the use of contingent valuation surveys to estimate non-use values for environmental goods, but do not claim to offer a full account of how to exclude non-self-interested preferences.¹⁹² This is indeed an important limitation both of restricting preferences accounts of welfare and of CBA.

187 J. Wayne Aldridge, Kent C. Berridge, & Terry E. Robinson, *Dissecting Components of Reward: 'Liking', 'Wanting', and Learning*, 9 CURRENT OPINIONS ON PHARMACOLOGY 65 (2009).

188 Yew-Kwang Ng, *Utility, Informed Preference, or Happiness: Following Harsanyi's Argument to Its Logical Conclusion*, 16 SOC. CHOICE WELFARE 197 (1999).

189 *Id.* at 210.

190 Bronsteen, Buccafusco, & Masur, *supra* note 18.

191 HAPPINESS & THE LAW, *supra* note 117, at 138–39.

192 ADLER & POSNER, *supra* note 1, at 39.

Bronsteen, Buccafusco, and Masur also point out that people may be mistaken about what outcomes will be best for them.¹⁹³ In one sense, this is uncontroversial. People frequently come to regret their choices. Sometimes, this may be because there is a conflict of interest between their present and past self, but often enough regret results from mistakes, as commonly understood. Therefore, Adler and Posner's account of welfare relies on fully informed preferences, even if they do not settle on a particular account of full information.¹⁹⁴ In another sense, even fully informed people may have preferences that do not reliably maximize their subjective well-being. In Bronsteen, Buccafusco, and Masur's account, these preferences are also mistakes that should be disregarded in favor of more reliable and direct indicators of welfare.¹⁹⁵ In Adler and Posner's view, one cannot be made better off by producing an outcome he never prefers and never would prefer, even if fully informed.¹⁹⁶ This class of cases seems to bottom out on a fundamental divergence of intuitions, with no clear resolution. In any case, Bronsteen, Buccafusco, and Masur succeed in casting doubt on the view that preference satisfaction, even in the restricted view outlined by Adler and Posner, offers a complete account of welfare. At the very least this should make policymakers more reluctant to defer to measures of individual preferences in marginal cases like those associated with pure time preference.

VIII. LAUNDERED PREFERENCES

Adler and Posner's specific account of welfare is Sophisticated Preferentialism, which holds that: "P is better off with S1, as compared to S2, just in case (1) P prefers S1 over S2; and (2) P would prefer S1 over S2 under ideal conditions; and (3) P's preference and ideal preference are suitably restricted."¹⁹⁷ Preference idealization under criterion (2) is designed to counter the objection that people's preferences may be "evil, ignorant, adaptive, or otherwise misshapen."¹⁹⁸ Adler and Posner refrain from committing themselves philosophically to either the full information, objectivist, or historical conception of preference idealization, but do claim that administrative agencies are better equipped to "launder" poorly informed preferences that are objectively bad (e.g., racist) or adaptive (e.g., a housewife who is indoctrinated to prefer subordination to her husband) preferences.¹⁹⁹ The upshot of criterion (3) is exclusion of preferences that are not self-interested, such as moral preferences.²⁰⁰ Adler and Posner readily acknowledge that appropriately defining the scope of self-interest is not straightforward, however.²⁰¹

Importantly, Adler and Posner concede that preference idealization, as they conceive it, does not solve the problem of conflicting preferences. They note that conflict

193 HAPPINESS & THE LAW, *supra* note 117, at 140.

194 ADLER & POSNER, *supra* note 1, at 38, 136–38.

195 HAPPINESS & THE LAW, *supra* note 117, at 140.

196 ADLER & POSNER, *supra* note 1, at 38, 138.

197 *Id.* at 36. See also Matthew D. Adler, *Beyond Efficiency and Procedure: A Welfarist Theory of Regulation*, 28 FLA. ST. L. REV. 241, 265 (2000).

198 ADLER & POSNER, *supra* note 1, at 33.

199 *Id.* at 38.

200 *Id.* at 39.

201 *Id.*

between preferences can be diachronic (i.e., change over time) or synchronic (e.g., a first order preference for watching TV over reading, but a second-order preference to be the kind of person who prefers to read).²⁰² These preference conflicts are precisely where we would expect time preference to come into play, suggesting that whatever the other merits of preference idealization, it does not speak to the central question of this article. However, time preference is potentially bound up with both objective idealization and the self-interestedness criterion. To see why, we must return to Parfit's defense of the rationality of pure time preference and his subsequent moral condemnation of it.

Consider the self-interestedness requirement. If we take the Parfitian idea of decreasing connectedness to successive future versions of one's self seriously, the concept of self-interest starts to break down. If saving for retirement is best viewed as mostly an idiosyncratic act of generosity toward a specific future person, then restricting preferences to self-interested ones becomes quite fraught when considering both the discount rate and mortality risks. Nonetheless, if we are willing to treat self-interestedness as a continuous scalar quantity rather than a stark binary, the self-interestedness requirement can be reconciled with Parfitian ideas about personal identity. If many diachronic preference conflicts are attributable to changes in feelings of connectedness to specific future selves, this could enable Adler and Posner's framework to better handle such cases. Many cases of synchronic preference conflicts can also be reconceptualized in terms of a conflict regarding how much to discount the future (watching TV might be more fun right now, but I can expect to reap benefits in the future if I choose to follow my second-order preference to read instead).

Now consider the requirement that preferences not be objectively bad. Parfit argues that even as we exonerate pure time preference from the charge of irrationality, we should replace *rational* condemnation of those who fail to take adequate account for the welfare of their future selves with *moral* condemnation.²⁰³ To the extent that pure time preference undervalues the interests of future people, we can condemn it as objectively bad within Adler and Posner's framework. What if we reject the Parfitian defense of the rationality of pure time preference, based on attenuated connectedness to future selves? Then pure time preference would be irrational, but not be immoral, since we might think people have a moral right to mistreat a future self they identify as much with as their current self without transgressing any moral principle. To handle this case, Adler and Posner's notion of objective badness would have to expand to include preferences that, though perhaps fully informed, are irrational. In either case, they should be willing to do so, given that their normative case for CBA depends on its status as "the welfare-maximizing decision procedure."²⁰⁴ Since pure time preference severs the connection between individual preferences and social welfare, any preference patterns that reflect it are unfit inputs for CBA.

202 *Id.* at 37.

203 PARFIT, *supra* note 27, at 170.

204 ADLER & POSNER, *supra* note 1, at 62.

IX. THE NUDGE DEBATE

Sunstein and Richard Thaler's *Nudge* thesis, also known as libertarian paternalism, prescribes crafting choice architecture to influence individual choices in normatively appealing directions consistent with enlightened preferences.²⁰⁵ Supporters of nudges should also favor circumscribing deference to individual preferences as articulated above. If public policy is justified in acting to influence self-regarding individual behaviors, then surely it should not defer to the preferences driving those behaviors when evaluating regulations targeted primarily at other-regarding activities. Moreover, I argue that even committed anti-paternalists who reject nudges can and should embrace my tweaks to CBA outlined above, which do not entail the same (arguable) autonomy infringements as nudges.

A common critique of Sunstein and Thaler's *Nudge* thesis attacks the distinction they draw between means and ends paternalism. Sunstein himself concedes that the distinction is sensitive to "the level of generality at which people's ends are to be described."²⁰⁶ At a sufficiently high level of generality, all paternalism can be characterized as ends paternalism. The distinction appears vulnerable in cases where individuals have multiple competing preferences, such as for bodily health on the one hand and unhealthy pleasures on the other. Aneil Kovvali points out that such cases are precisely the domain in which nudges operate.²⁰⁷ Kovvali offers a stylized example of a high school senior, Susan, choosing between three colleges she has applied to: Amherst, Bowdoin, and Carleton. Based on her research, Susan ranks the three colleges in the subject areas that interest her as follows:²⁰⁸

Economics	Political Science	History
1. Amherst	1. Bowdoin	1. Carleton
2. Bowdoin	2. Carleton	2. Amherst
3. Carleton	3. Amherst	3. Bowdoin

Susan then adopts a decision procedure under which, upon receiving two acceptance letters, she will consult her rankings and eliminate the school that wins in fewer categories. Kovvalli continues:

For example, if she receives acceptance letters from Amherst and Bowdoin, she will note that Amherst outperforms Bowdoin in Economics and History, while Bowdoin outperforms only in Political Science. As a result, she will discard the letter from Bowdoin.

She further resolves that if she receives a third letter, she will compare it to the surviving letter from the first two. So if she receives acceptance letters from Amherst and Bowdoin, and later Carleton, she will first discard the letter from

²⁰⁵ SUNSTEIN & THALER, *supra* note 5.

²⁰⁶ CASS R. SUNSTEIN, *WHY NUDGE?: THE POLITICS OF LIBERTARIAN PATERNALISM* (2014). *Why Nudge?* is based on Sunstein's Storrs Lecture at Yale Law School. See Cass R. Sunstein, *The Storrs Lectures: Behavioral Economics and Paternalism*, 122 *YALE L. J.* 1826 (2013).

²⁰⁷ Aneil Kovvali, *Who Are You Calling Irrational?*, 110 *Nw. U.L. REV.* 707, 712 (2016).

²⁰⁸ *Id.*

Bowdoin as described above. Once the offer from Carleton arrives, she will compare it to the surviving offer from Amherst and apply the same analysis. Noting that Carleton outperforms Amherst in Political Science and History, while Amherst outperforms only in Economics, she will discard the letter from Amherst.²⁰⁹

Kovvalli maintains that Susan's procedure is not irrational in the colloquial sense and that her preferences are entitled to respect.²¹⁰ Nonetheless, Susan's choice is sensitive to choice architecture:

If she receives letters from Amherst and Bowdoin alone, she will choose Amherst. If she receives letters from Bowdoin and Carleton alone, she will choose Bowdoin. But if she receives letters from Amherst and Carleton, she will choose Carleton. It follows that if she receives acceptance letters from all three colleges, her ultimate choice will depend on the order in which the letters are received: If she receives letters from Amherst, then Bowdoin, then Carleton, for example, she will choose Carleton; if she receives letters from Bowdoin, then Carleton, then Amherst, she will choose Amherst; if she receives letters from Amherst, then Carleton, then Bowdoin, she will choose Bowdoin; and so on. Susan's choice will depend on the context in which it is presented, even though the context conveys no relevant information.²¹¹

Kovvalli goes on to suggest that it would be illegitimate for Susan's father to manipulate Susan's decision by withholding the acceptance letter from his preferred option Amherst until after Susan has received acceptance letters from Carleton and Bowdoin and eliminated Carleton from consideration, leading her to ultimately select Amherst over Bowdoin. Kovvalli draws an analogy here to Sunstein and Thaler's "core cases" for nudging, where some choice architecture is inevitable. Even in such cases, Kovvalli maintains, nudges represent an exercise of power that favors one set of competing preferences over another.²¹²

Whether or not Kovvalli is right that Susan's preferences, as embodied in her subjective rankings and decision procedure, are entitled to respect in the sense of noninterference, it seems clear that the outcome of procedures like Susan's should not be respected in the sense of being used as an input in setting regulatory policy. Whatever can be said about the merits of Susan's decision procedure, it does not reveal any meaningful preference between the three colleges. On the information provided in Kovvalli's stylized example, Susan cannot be said to have any true preference among them. More information could enable a more meaningful decision. For instance, if Susan had information regarding the relative magnitudes of the differences in quality of the colleges in her areas of interest, or had any inclination with regard to which of the three subjects she is most likely to major in, or had any preferences and information on features of the colleges beyond those three subjects, she might have a meaningful preference. In the stylized hypothetical, the choice is perfectly symmetric.

Let us assume that Susan has researched thoroughly and this is the best information available to her at a reasonable search cost. In that case, a regulator that imputed a

209 *Id.* at 712–13.

210 *Id.* at 713.

211 *Id.*

212 *Id.* at 714–15.

revealed preference to Susan would be chasing a phantom. Whatever differences there are between the colleges in tuition, financial aid, location, campus culture, weather, alumni network, etc., are not factors into Susan's decision. Even regarding the three factors Susan does both care and have information about, it would be a mistake to infer that she cares most about economics if she chooses Amherst. That outcome would merely be a result of receiving the Amherst acceptance letter last or being rejected by Carleton, which outranks Amherst on two of the three factors.

One way of making sense of Susan's procedure is by reference to the multiple-selves model of the mind.²¹³ If we think of Susan as comprising three selves, each of whom *only* cares about the history, the political science, or the economics department, then information about the magnitudes of the differences in department quality would be irrelevant. From this perspective, Susan's decision procedure is analogous to a voting process with her three selves as the electors. Arrow's impossibility theorem holds that in cases with three or more distinct alternatives, no ordinal ranked voting system can ensure an outcome that meets a set of reasonable fairness conditions.²¹⁴ Indeed, given the preferences of Susan's selves, none of the three colleges is a Condorcet winner, meaning no college would win a majority vote against each of the other two colleges.²¹⁵ Since Susan must nonetheless pick a college, her decision procedure can be defended as a rational, if arbitrary, method of aggregating the preferences of her multiple selves. Since no procedure would ensure a fair outcome given the preference pattern of her selves, an external intervener seeking to nudge her decision would be acting based on interests outside the scope of Susan's preference. If this multiple-selves model were an accurate description of Susan's decision-making, it might sustain Kovvalli's critique of Sunstein and Thaler. However, once again, the arbitrary outcome of Susan's decision procedure would not be a solid basis for inferring any unitary preference on Susan's part. This severs any connection between any preference inferred from Susan's behavior and social welfare.

Given the specific features of Susan's case, one might expect the random contingencies of binary comparison and choice ordering to average out. In more realistic *Nudge* cases, like retirement savings or food choice, however, there are consistent patterns that would not average out. Such choices reflect the rate of time preference. Hyperbolic pure time preference is one way of explaining why some children "fail" the famous marshmallow test of delayed gratification.²¹⁶ Individuals with high rates of time preference will tend to save less for retirement and indulge more in unhealthy foods because they care less about future poverty or health problems than present consumption.

Sunstein and Thaler argue that these choices are sometimes mistakes and advocate non-coercive interventions—called nudges—to encourage more retirement savings and healthier food choices.²¹⁷ For instance, the Save More Tomorrow plan championed by Thaler enables people to commit in advance to contributing a higher percentage of

213 See David Lester, *A Multiple Self Theory of the Mind*, 1 COMPREHENSIVE PSYCH. 5 (2012).

214 See Kenneth J. Arrow, *A Difficulty in the Concept of Social Welfare*, 58 J. OF POL. ECON. 328 (1950).

215 See Duncan Black, *On the Rationale of Group Decision-making*, 56 J. OF POL. ECON. 23 (1948).

216 See Ebbe B. Ebbesen, Walter Mischel, & Antonette Raskoff Zeiss, *Cognitive and Attentional Mechanisms in Delay of Gratification*, 21 J. PERSONALITY & SOC. PSYCH. 204 (1972).

217 SUNSTEIN & THALER, *supra* note 5.

future pay increases to a retirement savings account.²¹⁸ This works because many people exhibit hyperbolic, rather than geometric, discounting, which can lead to preference reversal as the time approaches at which a sacrifice for a greater future reward is to be made approaches.²¹⁹ Likewise, Sunstein and Thaler advocate placing fruits and vegetables in school cafeterias at eye level and desserts in inconvenient locations.²²⁰ As applied to these cases, Kovvalli's critique would have to be that individuals have competing preferences for present consumption and for future health and wealth, and that any external intervention necessarily prioritizes one set of preferences rather than neutrally aiding individuals in fulfilling their own desires. As above, this critique, even if it is correct, offers no basis for policymakers to have confidence in relying on inferences of individual preferences in evaluating non-paternalistic regulations.

Other criticisms question the effectiveness of non-coercive nudges;²²¹ worry that nudges are infantilizing and undermine individuals' rational decision-making by depriving them of opportunities for practice;²²² and argue that covertly shaping individual choices may be a greater threat to liberty than open constraint.²²³ None of these criticisms apply to circumscribing the deference given to measures of individual preferences in CBA. Before policymakers move forward with more paternalistic interventions to correct arguably irrational individual behavior, they should stop deferring to the preferences driving those suspect behaviors for the purpose of evaluating non-paternalistic regulations.

X. OTHER ALTERNATIVES TO CBA

The foregoing has offered relatively mild criticism of the current practice of CBA and the potential of WBA. Perhaps the complications associated with reliance on measures of individual preference point to deeper problems with CBA. Indeed, several prominent scholars, including Frank Ackerman and Lisa Heinzerling, contend that CBA is fundamentally flawed and should be dispensed with entirely.²²⁴ A full assessment of their arguments is beyond the scope of this article. Instead, I will focus on two aspects of Ackerman and Heinzerling's critique that relate to the present inquiry: CBA requires applying monetary values to priceless human lives, and CBA trivializes the future.

Ackerman and Heinzerling's core contention about mortality risks is that every life is sacred and no amount of financial savings can justify allowing someone to die.²²⁵ They

218 Shlomo Benartzi & Richard H. Thaler, *Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving*, 112 J. POL. ECON. S164 (2004).

219 R.J. Herrnstein & Kris N. Kirby, *Preference Reversals Due to Myopic Discounting of Delayed Reward*, 6 PSYCH. SCI. 83 (1995).

220 SUNSTEIN & THALER, *supra* note 5, at 57.

221 Tom Goodwin, *Why We Should Reject 'Nudge'*, 32 POL. 85 (2012).

222 Karen Yeung, *Nudge as Fudge*, 75(1) MOD. L. REV. 122, 145 (2012).

223 Daniel M. Hausman & Brynn Welch, *Debate: To Nudge or Not to Nudge*, J. POL. PHIL. 123 (2010).

224 Ackerman & Heinzerling, *supra* note 17, at 1553. See also Martha C. Nussbaum, *The Costs of Tragedy: Some Moral Limits of Cost-Benefit Analysis*, 29(2) J. LEGAL STUD. 1005 (2000).

225 Ackerman & Heinzerling, *supra* note 17, at 1584.

disavow the interpretation that life has infinite value, claiming instead that “there is no ‘price’ for life because it’s value is immeasurable.”²²⁶ However, Ackerman and Heinzerling’s holistic alternative to attempting to measure the value of a life does not offer clear guidance for evaluating cases where regulators must select of precise level of stringency for regulating an environmental hazard that cannot be feasibly eliminated.²²⁷ For instance, how tightly to regulate the emission of a non-threshold pollutant. Ackerman and Heinzerling implicitly acknowledge that that society should not expend billions of dollars’ worth of resources to save a single life.²²⁸ But beyond references to a supposed consensus among “most system of ethical and religious belief,” they fail to offer much justification for the view that the value of a life is not quantifiable in principle.²²⁹

Instead, they focus their argumentative firepower on addressing the “standard response” of CBA defenders, that CBA only prices mortality risks, not identifiable human lives.²³⁰ For present purposes, we can grant the premise that Ackerman and Heinzerling succeed in collapsing the distinction between statistical and identifiable lives. This does not resolve the propriety of attaching a dollar value to human life, however, for two reasons. First, as Sunstein points out, there are tradeoffs between different mortality risks.²³¹ A strong precautionary approach to one risk may increase the total number of lost lives by magnifying another risk.²³²

A possible solution to this is risk-risk analysis, which seeks to minimize the overall mortality risk.²³³ Even this approach, however, cannot capture unanticipated future mortality risks that a less wealthy society will be less able to manage effectively. Second, even if policymakers could figure out how to balance known and unknown risks to minimize overall multi-generational mortality risks, people care about things other than not dying. It is true that that no amount of money, finite or otherwise, could compensate a person for the loss of his life, if delivered after his death.²³⁴ But people choose to accept avoidable mortality risks in the name of pleasure or convenience on a daily basis.²³⁵ They do not act as if avoiding death is lexically superior to all other values.²³⁶ This means they do not value their lives infinitely. But taking the idea that lives are sacred

226 FRANK ACKERMAN & LISA HEINZERLING, PRICELESS: ON KNOWING THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING 67 (2004) [hereinafter PRICELESS].

227 *Id.* at 214–15.

228 Ackerman & Heinzerling, *supra* note 17, at 1584.

229 PRICELESS, *supra* note 226. Ackerman and Heinzerling offer many compelling critiques of methodologies for valuing life. The validity of these methodologies is a distinct issue, however, from their claim that life cannot be valued with precision and in terms of units in terms of common units for the purposes of weighing against other goods (i.e., in dollars).

230 *Id.*

231 Cass R. Sunstein, *Health-Health Tradeoffs*, 63(4) CHICAGO L. REV. 1533 (1996).

232 Cass R. Sunstein, *Beyond the Precautionary Principle*, 151 U. PA. L. REV. 1003 (2003).

233 W. Kip Viscusi, *Risk-Risk Analysis*, 8 J. RISK & UNCERTAINTY 5 (1994).

234 PRICELESS, *supra* note 226, at 69. Though some people are willing to sacrifice their lives to improve the financial circumstances of their surviving family members.

235 Stephen Lyng, *Edgework: A Social Psychological Analysis of Voluntary Risk Taking*, 95(4) AM. J. SOC. 851 (1990).

236 JOHN RAWLS, A THEORY OF JUSTICE 37–38 (rev. ed. 1999). A lexically superior value is given absolute priority, such that it must be fully satisfied before considered other values. It is not subject to balancing against lexically inferior values.

and priceless seriously would mean refusing to make explicit tradeoffs between saving lives and other values. Fortunately, as indicated above, Ackerman and Heinzerling do not advocate assigning infinite value to human life.²³⁷ They merely require that policy-makers avoid specifying a finite value in terms of units that can be weighed directly against other social values.²³⁸

Instead, Ackerman and Heinzerling suggest relying on a menu of regulatory options that have emerged from experience, including “technology-based” regulation, pollution trading with caps set via political compromise, and information regulation.²³⁹ Information regulation, typically mandated disclosure, is not truly an alternative to CBA.²⁴⁰ It is a regulatory tool that may make sense in some contexts and not in others. CBA or another decision procedure must be used to decide whether and how much disclosure to mandate. It is true that the costs associated with information disclosure rules are typically too low to justify the decision costs associated with CBA,²⁴¹ but this does not make information regulation an alternative to CBA. Moreover, in cases where information regulation is used, a decision procedure is also needed to determine whether more robust measures are also justified.

Technology-based regulation and pollution trading, for their part, do not avoid tradeoffs between saving lives and other social objectives. At best, they hide those decisions or address them implicitly in legislative bargains. After all, technology-based standards typically do not eliminate deaths from any given form of pollution. Greater safety could be purchased with tougher standards that force the regulated industry to innovate, reduce output, or shut down entirely. Of course, this would have substantial social costs.

But if life is truly priceless, what basis do we have for declining to bear them? This is even more clear in the case of legislated targets. The legislative process may carry a badge of democratic legitimacy, but it is unlikely to produce standards that drive mortality risks down to their absolute minimum. This is both because regulated industries often have substantial political clout and because ordinary people are generally unwilling to sacrifice all other values for incremental increases in safety. Indeed, Ackerman and Heinzerling discuss a case where Congress hastily overturned a rule promulgated by the Occupational Safety and Health Administration that was justified under the agency’s CBA.²⁴² This would seem to be more an indictment of congressional judgment than of CBA, at least from Ackerman and Heinzerling’s pro-regulatory perspective. Likewise, Ackerman and Heinzerling reference Arrow’s Impossibility Theorem to undermine CBA’s claim to reflect a social welfare function, but Arrow’s proof applies equally to any decision procedure, including the democratic decision making they prefer.²⁴³

Whatever advantages technology-based regulation and legislated standards may have over CBA, they cannot avoid the inevitable tradeoffs between safety and other social goods. What they can do is avoid putting an explicit monetary price on the value of a life. But this is only an aesthetic advantage. In CBA, dollars are the common unit of

237 PRICELESS, *supra* note 226.

238 *Id.*

239 Ackerman & Heinzerling, *supra* note 17, at 1581–83.

240 *Id.*

241 LIVERMORE & REVESZ, *supra* note 8.

242 PRICELESS, *supra* note 226, at 106–07.

243 *Id.* at 209; Arrow, *supra* note 214.

social value used to allocate scarce resources to competing social goods.²⁴⁴ Declining to assess competing priorities in terms of a common unit of value only means that tradeoffs will be less transparent and less likely to maximize social welfare. As Adler and Posner point out:

[Technology-based regulation] is clearly suboptimal with regard to welfare. First, even if cutting-edge technology is optimal for some firms, it will not be optimal for all, given the heterogeneity of the welfare of firms' activities and the costs of employing the technology. Second, there is no reason to expect the process of technological development to generate a technology that is optimal on average. On the one hand, researchers motivated by prospects other than market demand, such as fame, may invent expensive technologies that firms or other actors (even if they did internalize all external effects and had good information) wouldn't buy. On the other hand, given external effects or poor information, technologies that result from market demand may fall below the welfare-maximizing level; a "technology forcing" policy might be better.²⁴⁵

Likewise, political bargaining over the level at which to set the cap for a pollution trading system *may* settle on the optimal amount of pollution, but Ackerman and Heinzerling offer no reason to expect this outcome. Certainly, their comparison between environmental protection and defense and counterterrorism spending does not inspire confidence.²⁴⁶ Nonetheless, the relative simplicity of emissions trading arguably offers a circumstance in which it is reasonable to think that legislation can and should specify the target, leaving the Environmental Protection Agency with minimal discretion. If so, CBA would indeed be inappropriate. One would still hope that legislative bargaining at least consider how much more or less mortality risk would be associated with a marginally higher or lower emissions cap and weigh risk reduction against other social goods. More generally, there are many domains of regulatory law and policy where the Congress lacks the capacity to legislate with a sufficient level of specificity such that agencies do not need a decision procedure beyond simply applying the statute.²⁴⁷ CBA, with all its flaws, is the best available procedure for agencies to balance conflicting social goods to maximize social welfare in these cases.

A similar analysis applies to Ackerman and Heinzerling's critique of the practice of discounting. As with mortality risk, agencies (or the Congress) must make some decision about how to balance the needs of the present and the future. One can question the merits of any choice, as this article has, but there is no way to avoid trading present welfare against future welfare. Perhaps this article does not go far enough in advocating for a lower discount rate.²⁴⁸ Certainly, some of the creative discounting practices criticized by Ackerman and Heinzerling are unjustifiable. For instance, discounting the value of lost life-years, the value of which was itself calculated from VSL estimates that already

244 ADLER & POSNER, *supra* note 1, at 91.

245 *Id.*

246 See PRICELESS, *supra* note 226, at 216–19.

247 See Harold H. Bruff, *Legislative Formality, Administrative Rationality*, 63 TEX. L. REV. 207 (1984).

248 See Cowen & Parfit, *supra* note 22 (arguing for a zero-discount rate, though taking account of many of its components directly).

reflect pure time preference, double-counts a preference that should only be counted once, if at all.²⁴⁹ Regardless, any regulatory regime will inevitably reflect some balance between the interests of the present and the future. Just as technology-based regulations and legislated emissions caps reflect an implicit valuation of human life, they reflect an implicit judgment regarding how heavily to weigh the interests of our future selves and future generations. If we try to avoid the question and refuse to explicitly quantify the balance we strike, we will only succeed in addressing the tradeoff in an opaque and inconsistent manner.

Finally, much of Ackerman and Heinzerling's critique targets the use of CBA as a tool to justify deregulation. Many of their criticisms of implementations of CBA are compelling. However, much of this discussion cuts against Ackerman and Heinzerling's contention that CBA is *inherently* biased against regulation, rather than merely susceptible to being wielded by policy analysts with an anti-regulatory agenda.²⁵⁰ Although I disagree with some of their claims, Revesz and Livermore's *Retaking Rationality* and the subsequent work of the Institute for Policy Integrity constitute an existence proof for the use of CBA to advance a pro-regulatory agenda. Similarly, Viscusi points out that refusing to apply a monetary value to mortality risks tends to produce *less* protective regulatory standards.²⁵¹ Ackerman and Heinzerling are right to point out that CBA is not inherently neutral and is subject to manipulation by anti-regulatory advocates and scholars. However, as Revesz and Livermore argue, the correct response is to offer a positive vision of cost-benefit analysis that seeks to assess the true costs and benefits of regulations in good faith.²⁵² This article is a contribution to that effort.

XI. CONCLUSION

Neither current restricted-preference accounts nor happiness survey-based approaches appear to fully capture social welfare. Adler and Posner offer compelling examples of cases where happiness is unaffected by some difference in the state of the world, but nonetheless that difference seems to matter. However, there are also many cases when even preferences that are informed and survive the forms of idealization endorsed by Adler and Posner point to non-optimal outcomes. In the clearest case, relating to pure time preference and discount rates, presently existing people are imperfect proxies for their future selves.

Whether this results from a failure of rationality or of ethics, policymakers should decline to ratify this preference pattern. The case of mortality risk is more complicated, but the preference patterns that are used to justify the VSL method similarly fail on normative grounds. While treating all life-years as equally valuable is probably not the optimal approach, it is a clear improvement over treating all deaths as equally bad.

More generally, policymakers should be more open to the conclusion that certain sorts of preferences should not be deferred to. At a minimum, any domain in which paternalistic policies designed to alter individual choices are considered, even if those

249 PRICELESS, *supra* note 226, at 196–97.

250 LIVERMORE & REVESZ, *supra* note 8.

251 VISCUSI, *supra* note 93, at 9.

252 LIVERMORE & REVESZ, *supra* note 8, at 31, 42–45.

interventions are noncoercive nudges, should give policymakers pause about evaluating regulations addressing other-regarding activities based on the preferences driving those behaviors. Even a committed opponent of all forms of paternalism need not conclude that deference to individual preferences should be absolute in the context of non-paternalistic policymaking. In addition to the restrictions endorsed by Adler and Posner, preferences must also have a plausible rational justification that maintains a connection to social welfare. This is a lax standard designed to prevent expanding the discretion of regulators unnecessarily, but nonetheless rules out a number of preference patterns currently relied upon.

However, caution is warranted in excluding preferences on the ground of irrationality. For any pattern of preferences—including intransitive preferences, preferences influenced by framing effects or irrelevant alternatives, etc.—there is generally some plausible rational justification. This justification may appeal to a multiples-selves model, to a heuristic that may be rational to adopt as a general rule given information costs, or to some form of meta-rationality (e.g., anger as a strategic pre-commitment device for costly punishment).²⁵³ After all, if a preference had no rational basis whatsoever, one might expect natural selection to have weeded it out.²⁵⁴ This evolutionary logic may cut both ways, however. Perhaps certain preferences were rational in the evolutionary environment, but no longer are.²⁵⁵ Perhaps other irrational preferences are linked via gene complexes to other traits that provide enough fitness to allow the irrationality to persist.

We need not resolve this debate over the existence of irrational preferences. Instead, in considering whether a preference is a proper input for policy analysis, we should consider whether any of the plausible rational justifications maintain the connection between preferences and social welfare. In cases like Susan's procedure for picking among her college options, any preference that might be inferred from her choice carries no information of relevance to social welfare maximization. In other cases, like how to weigh the interests of foreigners, this judgment may be less clear. There may be a strong normative case for a more cosmopolitan approach to policy analysis, but this is a sufficiently contested ethical issue that we cannot really say that most people's preference for aiding their fellow citizens carries no normative weight. Indeed, the key question in this regard is whose welfare public policy should seek to maximize.

The criterion that I have advocated for determining the relevance of individual preferences assumes we know whose welfare public policy should serve. Neutrality between domestic citizens is well-established, but there is no equivalent normative consensus regarding how heavily public policy should weigh the welfare of foreigners. This article does not seek to fully resolve the debate about preferences and public policy, but to offer a framework for evaluating when we should consider measures of individual preferences

253 See generally ROBERT H. FRANK, *PASSIONS WITHIN REASON: THE STRATEGIC ROLE OF THE EMOTIONS* (1988).

254 See Alan R. Rogers, *Evolution of Time Preference by Natural Selection*, 84 AM. ECON. REV. 460 (1994); Thomas Grund, Dirk Helbing, & Christian Waloszek, *How Natural Selection Can Create Both Self- and Other-Regarding Preferences, and Networked Minds*, 3 SCI. REP. 1480 (2013); Marcus Salomonsson & Jorgen W. Weibull, *Natural selection and social preferences*, 239 J. THEORETICAL BIOLOGY 79 (2006).

255 See John D. Balling & John H. Falk, *Evolutionary Influence on Human Landscape Preference*, 42 ENV'T & BEHAV. 479 (2009).

as proper inputs for policy evaluation. In important cases, like discounting and valuing life, it offers clear guidance. In other contexts, disputes will remain. Even in these cases, articulating disputes in terms of the framework advanced in this article should enable a more productive debate.

Gabriel Weil is a Senior Research Associate at the Climate Leadership Council and a Visiting Researcher at Georgetown University Law Center. The views expressed in this article are his own and do not necessarily reflect those of the Climate Leadership Council or its staff. The author would like to thank Richard Revesz for a conversation that inspired this article; Arden Rowell, Alejandro Camacho, John Bronsteen, Jonathan Masur, Matthew Adler, Cass Sunstein, Julie Zauzmer, and Todd Rubin for their invaluable feedback on early drafts, and the staff of the Texas Environmental Law Journal and its external reviewers from the Texas State Bar Environmental & Natural Resources Law section for their excellent editing.

