Optimal Pension Plan Default Policies when Employees are Biased

Asen Ivanov

Working paper No. 893

September 2019

ISSN1473-0278

School of Economics and Finance

Queen Mary
University of London

Optimal Pension Plan Default Policies when Employees are Biased

Asen Ivanov*

Queen Mary University of London

September 9, 2019

Abstract

What is the optimal default contribution rate or default asset allocation in pension plans? Could active decision (i.e., not setting a default and forcing employees to make a decision) be optimal? These questions are studied in a model in which each employee is biased regarding her optimal contribution rate or asset allocation. In this model, active decision is never optimal and the optimal default is, depending on parameter values, one of three defaults. The paper also explores how the parameters affect the optimal default and the total loss in the population at the optimal default.

Keywords: optimal defaults, libertarian paternalism, nudging, pension plan design

JEL: D14, D91, J26, J32

*School of Economics and Finance, 327 Mile End Road, London, United Kingdom; tel. +44 (0)20 7882 5886; a.ivanov@qmul.ac.uk

1 Introduction

A large literature documents that the default contribution rate and the default asset allocation in defined contribution pension plans have a large affect on employees' saving for retirement.[1] This raises the question of what is the optimal default policy. This question can be broken down into two parts. First, what is the optimal default? Second, how does the optimal default compare to "active decision" (AD), i.e., to not setting a default and forcing employees to make a decision?

Although these questions have been studied previously, most of the existing literature on optimal default policies assumes either that individuals know their optimal options or that they have unbiased beliefs about them.[2] This assumption runs counter to the common view that, left to their own devices, individuals would make systematic mistakes (e.g., by undersaving) and defaults can be used to "nudge" them towards better options.[3]

The current paper analyses a model in which employees are biased. In particular, the model has the following key features. First, for each employee, there is an optimal option, x, which corresponds to her optimal contribution rate or to the optimal fraction of her pension plan portfolio invested in stocks (with the rest of the portfolio being invested in risk-free bonds).

Second, each employee is biased in the sense that, if her optimal option is x, her preferred option (i.e., the one she would choose) is $x - b$. One can interpret the bias, b, as a mistake. In the context of contribution rates, this interpretation should be appealing for proponents of the view that individuals undersave for retirement from the

[1] See Madrian and Shea (2001), Choi et al. (2002, 2003, 2004a, 2004b, 2006), Beshears et al. (2008), Carroll et al. (2009), and Chetty et al. (2014). Bronchetti et al. (2013) observe no default effect.

[2] The only exceptions that I am aware of, Goldin and Reck (2018) and Ivanov (2019), are discussed in section 6.

[3] For example, see Thaler and Sunstein (2003).

point of view of their own interest.[4] In the context of asset allocation, it is plausible that young people underweight stocks because (i) as a result of exponential growth bias,[5] they underestimate the long-term difference in wealth that results from investing in stocks at a, say, 7% annual return rather than in bonds at a, say, 1% annual return or (ii) they are not aware of the high historical equity premium[6] or of the fact that historically equity returns have dominated bond returns over long horizons.[7,8] Alternatively, one can view the bias as reflecting an externality rather than a mistake. For example, if a perfectly rational employee saves little for retirement because she anticipates being bailed out by the government, she would not be undersaving from the point of view of her own interest, but would be undersaving from the point of view of the government.

Third, if an employee has to decide actively, either because she opts out of the default or because the default policy is AD, she incurs a cost, c.

Fourth, employees are heterogeneous–their optimal options are assumed to be distributed (for much of the paper, uniformly distributed) on an interval $[\underline{x}, \overline{x}]$ of length 2ϵ. The parameter ϵ will capture the degree of heterogeneity in the population.

Finally, if an employee ends up with option x' when her optimal option is x, this

[4]There is currently a heated theoretical and empirical debate on whether individuals do indeed undersave for retirement. For an overview of the issues as well as further references, see Poterba (2015) as well as section 4.6 in Ericson and Laibson (2019).

[5]For evidence of exponential growth bias, see Eisenstein and Hoch (2005), Stango and Zinman (2009), as well as Levy and Tasoff (2016).

[6]There is some suggestive evidence of such unawareness: Beshears et al. (2017) report that experimental subjects invest more heavily in stocks when they are shown data on historical returns.

[7]Benartzi and Thaler (1999) report that experimental subjects invest more heavily in stocks when shown 30-year rather than 1-year return distributions. Beshears et al. (2017) find that this effect is not robust, at least when subjects are shown 5-year vs. 1-year return distributions.

[8]There is evidence that education (Campbell (2006) and Calvet, Campbell, and Sodini (2007)) and IQ (Grinblatt, Keloharju, and Linnainmaa (2011)) have a positive effect on stock market participation. This is consistent with the view that nonparticipation, or, more generally, low investment in equities, is often a mistake. Bach, Calvet, and Sodini (2016) show that wealthier households in Sweden hold riskier financial portfolios. The associated higher average returns substantially increase inequality in financial wealth. Thus, to the extent that limited exposure to stocks is a mistake, it is one that contributes to inequality.

entails a loss from deviations equal to $l(x-x')$. For much of the paper, we will assume that $l(\cdot)$ is strictly convex and symmetric around zero. Each employee's loss equals her loss from deviations plus, possibly, the cost c. An optimal default policy is one that minimises the total loss in the population.

After presenting the details of this model in section 2, we turn to the main findings in sections 3-5. The first main finding is that AD is never an optimal default policy.

The second main finding is that, given the values of the parameters b, c, and ϵ, the optimal default is one of three defaults. The first default is placed in the middle of $[\underline{x}, \overline{x}]$. We call this the centre (C) default. If all employees would stay with the C default, it is optimal.

The second default is such that employees with $x = \underline{x}$ think it's too high, but are only just willing to stay with it. We call this the $\underline{x}$-perceived-equal-loss-from-in-or-out ($\underline{x}$-InOut-PEL) default. This default is optimal when (i) some employees would opt out of the C default and (ii) an employee with $x = \overline{x}$ is close enough to the $\underline{x}$-InOut-PEL default that staying with it involves a loss lower than the loss from opting out. When the $\underline{x}$-InOut-PEL default is optimal, all employees stay with it.

The third default is such that it is below $\overline{x}$ and employees with $x = \overline{x}$ incur a loss from staying equal to the loss they would incur from opting out. We call this the $\overline{x}$-equal-loss-from-in-or-out ($\overline{x}$-InOut-EL) default. This default is optimal when neither of the other two types of defaults are optimal. When this default is optimal, employees with x's above a cutoff stay while employees with x's below the cutoff opt out.

The third main finding is about how each parameter affects the optimal default. Loosely speaking, increasing b or ϵ tilts the planner away from the C default and towards the $\overline{x}$-InOut-EL default, whereas increasing c has the opposite effect.

The fourth main finding is that the total loss in the population at the optimal

default is weakly increasing in b, either weakly decreasing or nonmonotone in c, and strictly increasing in ϵ. As a result, if the planner can lower employees' bias or decrease heterogeneity in the targeted population, doing so is beneficial (or at least doesn't hurt); if the planner can affect the cost c, she may want to either increase it or decrease it depending on specifics.

For purposes of tractability, the model is designed to be as simple as possible. As a result, it lacks realism in many ways. Section 6 discusses some of the model's assumptions as well as potentially relevant factors that are omitted from it. Because the existing literature on optimal defaults deals with some of these factors, section 6 will also serves as a literature review. Section 7 concludes.

2 The Model

The model has the following components. First, for each employee, there is an optimal option, x, which corresponds either to her optimal contribution rate or to the optimal fraction of her pension plan portfolio invested in stocks (with the rest of the portfolio being invested in risk-free bonds). Employees' optimal options are distributed on $[\underline{x}, \overline{x}] \subseteq [0, 1]$, where $\underline{x} < \overline{x}$. We let 2ϵ denote the width of this interval.

Second, each employee is biased in the sense that, if her optimal option is x, her preferred option is $x - b$.[9] We assume $b > 0$, which is without loss of generality.[10] Note that, if $b > \underline{x}$, then $x - b < 0$ for some employees. Allowing some employees to choose $x - b < 0$ is unrealistic given that employees typically cannot choose a negative contribution rate or to short-sell stocks in their retirement portfolios. To avoid this

[9] An additive bias simplifies the analysis. However, there is no reason to believe that an additive bias is more realistic than, say, a multiplicative bias.

[10] If each employee sets the contribution rate or the allocation to stocks too high, we can think of x as the optimal fraction of her salary that she does not contribute towards her pension plan or as the optimal fraction of her portfolio invested in bonds, respectively. With this relabelling, the assumption $b > 0$ would be appropriate.

possibility, we assume $b \leq \underline{x}$.

Third, if an employee has to decide actively, either because she opts out of the default or because the default policy is AD, she incurs a cost, $c > 0$. We refer to c as the opt-out cost, even though there is no opting out under AD. We interpret c as reflecting implementation costs, i.e., the time and effort required to contact the relevant people in the human resources department, to fill in the necessary paperwork, etc.

Fourth, if an employee ends up with option x' when her optimal option is x, we assume this entails a loss from deviations equal to $l(x - x')$ and a perceived loss from deviations (from the point of view of the employee) equal to $l(x-b-x')$. We maintain throughout the following assumptions on $l(\cdot)$: (i) $l(\cdot)$ is continuous, (ii) $l(0) = 0$, (iii) $l(\cdot)$ is strictly decreasing on $(-\infty, 0]$ and strictly increasing on $[0, \infty)$, and (iv) $\lim_{z\to-\infty} l(z) = \lim_{z\to\infty} l(z) = \infty$. These assumptions should be uncontroversial. Later, we will impose additional assumptions on $l(\cdot)$.

In this setup, how does an employee with optimal option x behave and what loss does she incur? Under AD, the employee incurs the cost c and chooses $x - b$. As a result, she incurs a loss equal to $c + l(x - (x - b)) = c + l(b)$.

How about if she faces a default, D? Letting $\Delta = x - D$, the loss associated with staying with the default is $l(\Delta)$. However, the perceived loss is $l(x-b-D) = l(\Delta-b)$. The employee stays with the default if and only if $c \geq l(\Delta - b)$. The latter inequality can be written as $\Delta_L \leq \Delta \leq \Delta_R$, where Δ_L and Δ_R are the two values of Δ that solve the equation $c = l(\Delta - b)$ (see Figure 1). Note that, given that an employee with $\Delta = b$ thinks the default is ideal, an employee with $\Delta = \Delta_L/\Delta = \Delta_R$ thinks the default is too high/low but is only just willing to stay with it. If an employee opts out, she incurs the cost c and chooses $x-b$, so that her loss is the same as under AD, namely $c + l(b)$.

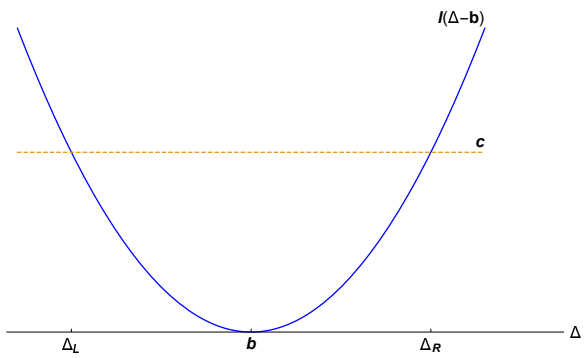

Figure 1: Determination of Δ_L and Δ_R.

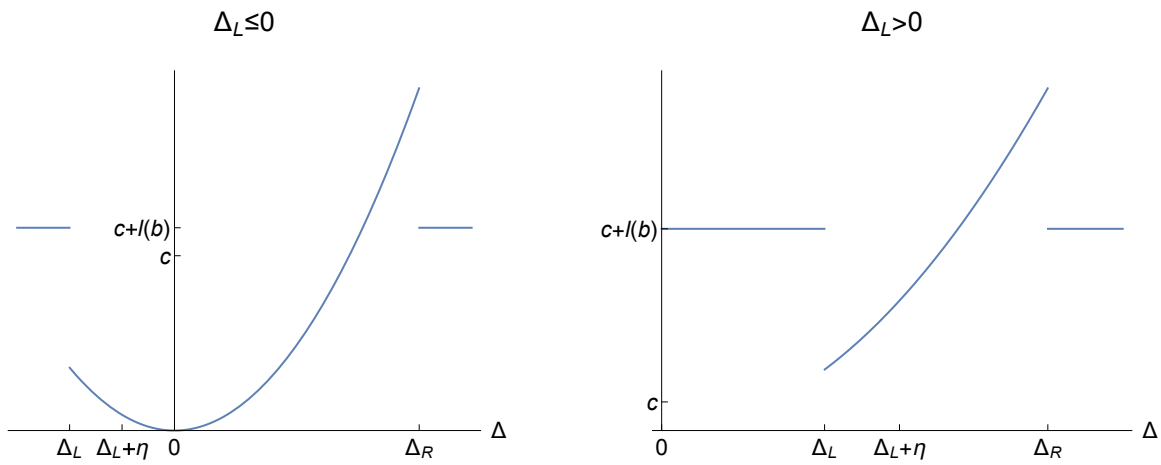

Figure 2: Graph of $L(\cdot)$.

Given the remarks in the previous paragraph, the loss of an employee with optimal option x who faces a default D is captured by the following function:

$$L(\Delta) = \begin{cases} l(\Delta) & \text{if } \Delta_L \leq \Delta \leq \Delta_R \\ c + l(b) & \text{otherwise} \end{cases}. \tag{1}$$

Figure 2 shows the graph of $L(\cdot)$ for each of the cases $\Delta_L \leq 0$ and $\Delta_L > 0$. The following lemma establishes that key features of Figure 2 hold more generally.[11]

Lemma 1

[11] All proofs are in the appendix.

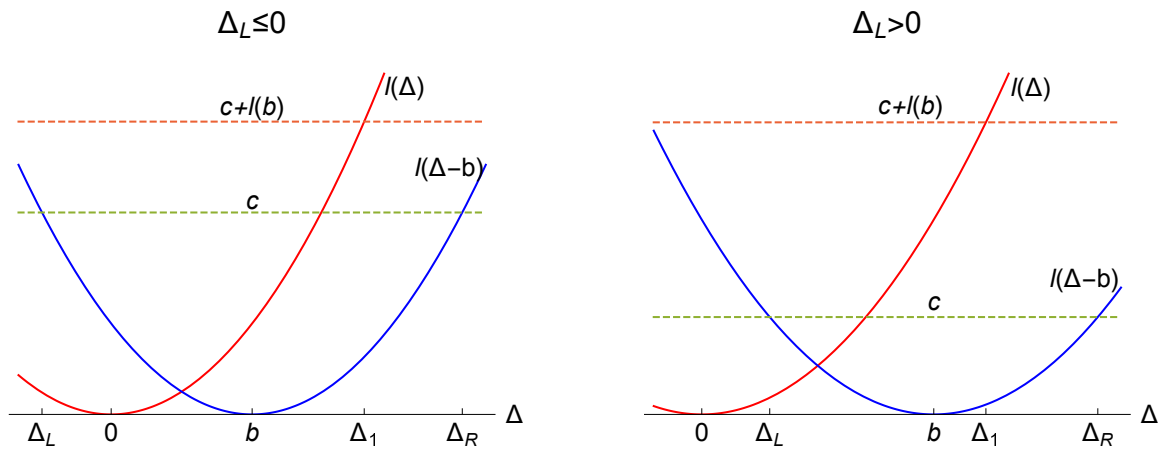

Figure 3: Actual loss, $l(\Delta)$, and perceived loss, $l(\Delta - b)$, from staying with the default.

1) If $\Delta_L \leq 0$, $L(\Delta_L) < c$.

2) If $\Delta_L > 0$, $L(\Delta_L) < l(b)$.

3) If $l(\cdot)$ *is strictly convex,* $L(\Delta_R) > c + l(b)$.

The logic behind statement 1) goes as follows. Assume $\Delta_L \leq 0$ and consider an employee with $\Delta = \Delta_L$. The default is too high for this employee ($\Delta \leq 0$ is equivalent to $x \leq D$), and the employee recognises this (recall that an employee with $\Delta = \Delta_L$ thinks the default is too high). However, given her bias, she exaggerates the extent to which the default is too high and, as a result, exaggerates the loss from staying with it. Given that an employee with $\Delta = \Delta_L$ thinks that the loss from staying with the default is c, it must be that the true loss from staying, $l(\Delta_L)$, is less than c (see the left panel in Figure 3).

Statement 2) holds because, when $\Delta_L > 0$, an employee with $\Delta = \Delta_L$ is closer to the default than an employee with $\Delta = b$, so that $l(\Delta_L) < l(b)$ (see the right panel in Figure 3). Given that $L(\Delta_L) = l(\Delta_L)$, the latter inequality is equivalent to $L(\Delta_L) < l(b)$.

The argument behind statement 3) is the following. An employee with $\Delta_R - b$ faces a loss of c from staying with the default (by the definition of Δ_R). Given that $\Delta_R - b > 0$ and $l(\cdot)$ is strictly convex, the additional loss from staying for an employee with $\Delta = \Delta_R$, $l(\Delta_R) - l(\Delta_R - b)$, is larger than $l(b) - l(0) = l(b)$.

Statements 1) and 2) in Lemma 1 imply that $L(\Delta_L) < c + l(b)$, so that $L(\cdot)$ has a downward jump at Δ_L. Statement 3) says that $L(\cdot)$ has a downward jump at Δ_R. The downward jumps of $L(\cdot)$ at Δ_L and Δ_R will play a key role in determining the optimal default policy.

3 The Optimal Default Policy

Let us observe that, given that employees' x's lie in $[\underline{x}, \overline{x}]$, employees' corresponding Δ's given a default D lie in $[\underline{x} - D, \overline{x} - D]$. Thus, by choosing D the planner is effectively shifting around the latter interval on the horizontal axis in each panel in Figure 2. Thus, denoting the upper endpoint of this interval as $\bar{\Delta}$ (i.e., $\bar{\Delta} = \overline{x} - D$), we can think of the planner as directly choosing $\bar{\Delta}$ rather than D. In fact, this turns out to be notationally more convenient. Moreover, given any $\bar{\Delta}$, it is easy to infer the position of the corresponding default, D, relative to $[\underline{x}, \overline{x}]$. In particular, the position of D relative to $[\underline{x}, \overline{x}]$ is the same as the position of $\Delta = 0$ relative to $[\bar{\Delta} - 2\epsilon, \bar{\Delta}]$. E.g., if $\Delta = 0$ is below/in the lower end of/in the middle of/in the upper end of/above $[\bar{\Delta} - 2\epsilon, \bar{\Delta}]$, then D is below/in the lower end of/in the middle of/in the upper end of/above $[\underline{x}, \overline{x}]$.

3.1 AD

Before exploring how the planner optimally sets $\bar{\Delta}$, let us consider whether AD could ever be an optimal default policy.

Proposition 1 *AD is never an optimal default policy.*

The logic for Proposition 1 is the following. By the fact that $L(\Delta_L) < c + l(b)$ and the continuity of $l(\cdot)$, there exists $\eta > 0$, such that $L(\Delta_L + \eta) < c + l(b)$. If the planner sets $\bar{\Delta} = \Delta_L + \eta$, then (i) no employee incurs a loss strictly greater than $c + l(b)$ because $L(\Delta) \leq c + l(b)$ for all $\Delta \leq \Delta_L + \eta$ (see Figure 2) and (ii) employees with Δ's between Δ_L and $\Delta_L + \eta$ incur a loss strictly less than $c + l(b)$ (again, see Figure 2). Thus, the total loss in the population with $\bar{\Delta} = \Delta_L + \eta$ is strictly lower than the total loss with AD.

3.2 The Optimal Default

From here on we maintain the following assumptions in addition to those imposed in section 2: employees' optimal options are *uniformly* distributed on $[\underline{x}, \overline{x}] \subseteq [0, 1]$, $l(\cdot)$ is strictly convex, and $l(\cdot)$ is symmetric around 0 (i.e., $l(z) = l(-z)$ for all z). These assumptions are discussed in section 6.

Given Proposition 1, an optimal default is an optimal default policy. We now turn our attention to the question of how the planner sets the default, or equivalently $\bar{\Delta}$, optimally.

Observe that the total loss in the population associated with a given value of $\bar{\Delta}$ equals $\frac{1}{2\epsilon}$ times the area under $L(\cdot)$ over the interval $[\bar{\Delta} - 2\epsilon, \bar{\Delta}]$. An optimal $\bar{\Delta}$ solves:

$$\min_{\overline{x}-1 \leq \bar{\Delta} \leq \overline{x}} \frac{1}{2\epsilon} \int_{\bar{\Delta}-2\epsilon}^{\bar{\Delta}} L(\Delta) d\Delta. \tag{2}$$

The constraint is equivalent to $0 \leq D \leq 1$, which reflects the natural constraint that the planner cannot set a default contribution rate or a default allocation to stocks outside of $[0, 1]$.

The following proposition characterises the solution to problem (2).

Proposition 2 *Let Δ_1 denote the value of $\Delta > 0$, such that $L(\Delta) = c + l(b)$ (see Figure 3). The unique solution to problem (2) is:*

$$\bar{\Delta}^*(b, c, \epsilon) = \begin{cases} \epsilon & \text{if } \Delta_L \leq -\epsilon \\ \Delta_L + 2\epsilon & \text{if } \Delta_L > -\epsilon, \Delta_1 - \Delta_L \geq 2\epsilon \\ \Delta_1 & \text{if } \Delta_1 - \Delta_L < 2\epsilon \end{cases} . \tag{3}$$

The first case in expression (3) is illustrated in Figure 4. The shaded area in the figure represents the total loss in the population (scaled by 2ϵ). $\bar{\Delta} = \epsilon$ corresponds to the C default. As is evident from the figure, nobody opts out.

To see the intuition for why $\bar{\Delta} = \epsilon$ solves problem (2) when $\Delta_L \leq -\epsilon$, consider shifting $\bar{\Delta}$ slightly to the right or left away from ϵ in Figure 4.[12] If we shift $\bar{\Delta}$ slightly to the right by a small number $\delta > 0$, the shaded area increases by the area under $L(\cdot)$ over the interval $[\epsilon, \epsilon + \delta]$ and decreases by the area under $L(\cdot)$ over the interval $[-\epsilon, -\epsilon + \delta]$. Because $L(\cdot)$ has higher values just to the right of ϵ than just to the right of $-\epsilon$, the shift to the right leads to a net increase in the shaded area and is thus undesirable. A similar logic reveals that shifting $\bar{\Delta}$ slightly to the left is undesirable. An analogous argument involving small shifts to the right or left of $\bar{\Delta}^*(b, c, \epsilon)$ will apply to the second and third cases in expression (3), so that we don't repeat the argument when discussing these cases below.

The second case in expression (3) is illustrated in the two panels of Figure 5 (one panel for $\Delta_L \leq 0$ and one for $\Delta_L > 0$). $\bar{\Delta} = \Delta_L + 2\epsilon$ corresponds to the $\underline{x}$-InOut-PEL default. As is evident from the figure, nobody opts out. Note that employees with low x's (i.e., with Δ's near Δ_L) incur a smaller loss than those with high x's (i.e., with Δ's near $\Delta_L + 2\epsilon$). However, increasing the default to better accommodate employees with high x's (by lowering $\bar{\Delta}$ below $\Delta_L + 2\epsilon$) is undesirable because it

[12] The argument given here merely shows that $\bar{\Delta} = \epsilon$ is a local minimiser. The formal proof in the appendix establishes that it is also a global minimiser.

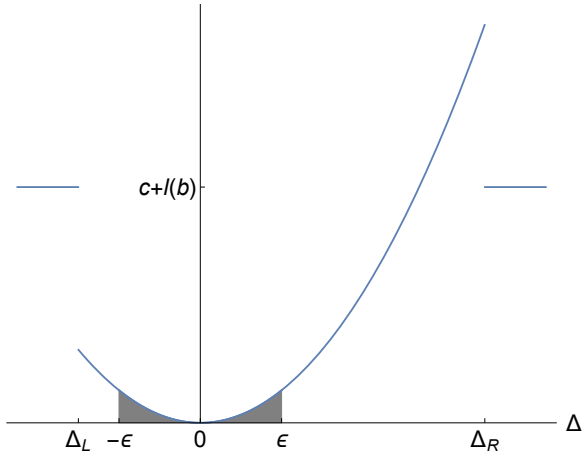

Figure 4: $\bar{\Delta}^*(b, c, \epsilon) = \epsilon$ (C default).

causes employees with low x's to opt out, which leads to even higher losses.

The third case in expression (3) is illustrated in the two panels of Figure 6 (one panel for $\Delta_L \leq 0$ and one for $\Delta_L > 0$).[13] $\bar{\Delta} = \Delta_1$ corresponds to the $\overline{x}$-InOut-EL default. As is evident from the figure, employees with x's above a cutoff (i.e., with Δ's above Δ_L) stay while employees with x's below the cutoff (i.e., with Δ's below Δ_L) opt out. Note that the planner could induce more employees to stay with the default if she slightly lowers it, i.e., if she shifts $\bar{\Delta}$ slightly to the right of Δ_1.[14] However, this would lead employees with high x's to incur losses above $c + l(b)$, which would more than undo any gains from getting more employees with lower x's to stay.

[13]The right panel is drawn so that $\Delta_1 - 2\epsilon > 0$. However, we could just as well have $\Delta_1 - 2\epsilon \leq 0$.

[14]The proof of Proposition 2 shows that $\Delta_1 < \overline{x}$ in the third case in expression (3), so that shifting $\bar{\Delta}$ slightly to the right of Δ_1 won't violate the constraints in problem (2).

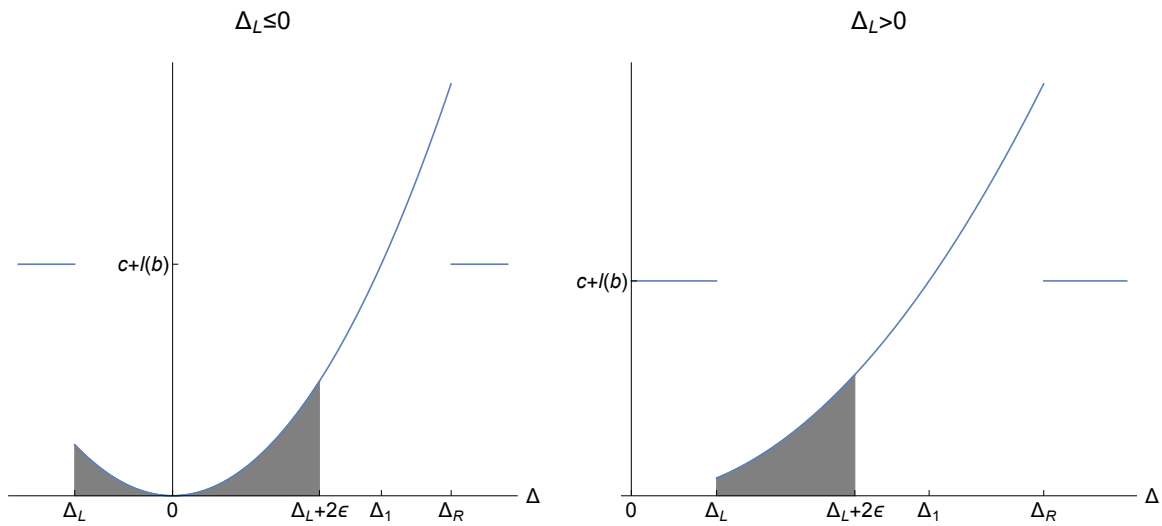

Figure 5: $\bar{\Delta}^*(b, c, \epsilon) = \Delta_L + 2\epsilon$ ($\underline{x}$-InOut-PEL default).

4 How Does Each Parameter Affect the Optimal Default?

In this section, we explore how each parameter affects which of the three cases in expression (3) holds and, hence, whether the C, $\underline{x}$-InOut-PEL, or $\overline{x}$-InOut-EL default is the optimal one.

Let us start with the following lemma.

Lemma 2 *$\Delta_1 - \Delta_L$ is strictly decreasing in b and strictly increasing in c.*

The logic for why $\Delta_1 - \Delta_L$ is strictly decreasing in b is the following. Increasing b by $\sigma > 0$ shifts the $c+l(b)$ line in each panel of Figure 3 up by $l(b+\sigma)-l(b)$. Given the fact that $\Delta_1 > b$[15] and the strict convexity of $l(\cdot)$, $l(\Delta_1+\sigma)-l(\Delta_1) > l(b+\sigma)-l(b)$, so that Δ_1 needs to increase by less than σ to accommodate the increase in b. Thus, Δ_1 increases less than one-for-one with b whereas Δ_L increases one-for-one with b (see Figure 1).

[15] $\Delta_1 > 0$ and $l(\Delta_1) = c + l(b) > l(b)$ imply $\Delta_1 > b$.

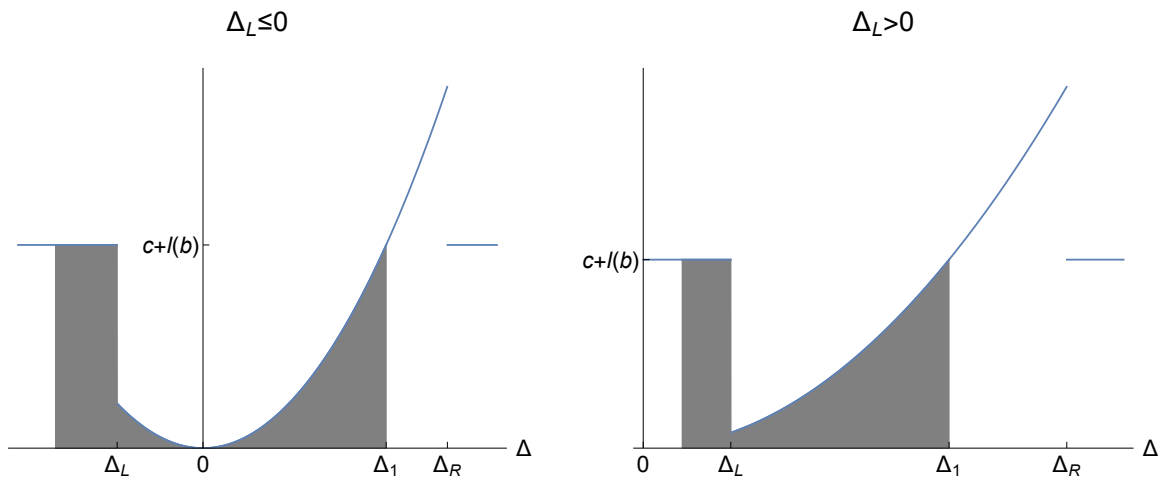

Figure 6: $\bar{\Delta}^*(b, c, \epsilon) = \Delta_1$ ($\overline{x}$-InOut-EL default).

That $\Delta_1 - \Delta_L$ is strictly increasing in c follows from the fact that Δ_1 is strictly increasing in c and the fact that Δ_L is strictly decreasing in c (see Figure 3).

We are ready to state the main result in this section.

Proposition 3

1) *Increasing b contracts/expands the set of (c, ϵ)-pairs for which the C/$\overline{x}$-InOut-EL default is optimal.*

2) *Increasing c expands/contracts the set of (b, ϵ)-pairs for which the C/$\overline{x}$-InOut-EL default is optimal.*

3) *Increasing ϵ contracts/expands the set of (b, c)-pairs for which the C/$\overline{x}$-InOut-EL default is optimal.*[16]

Statement 1) follows because increasing b (i) increases Δ_L (see Figure 1), thus making the inequality in the first case in expression (3) harder to satisfy, and (ii) by Lemma 2, decreases $\Delta_1 - \Delta_L$, thus making the inequality in the third case in expression (3) easier to satisfy.

[16]The terms "contracts" and "expands" are used in the sense of strict set inclusion.

Statement 2) holds because increasing c (i) decreases Δ_L (see Figure 1), thus making the inequality in the first case in expression (3) easier to satisfy, and (ii) by Lemma 2, increases $\Delta_1 - \Delta_L$, thus making the inequality in the third case in expression (3) harder to satisfy.

Statement 3) follows because increasing ϵ makes the inequality in the first/third case in expression (3) harder/easier to satisfy.[17,18]

5 How Does Each Parameter Affect the Total Loss?

How does each parameter affect the total loss in the population at the optimal default, $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$? The following proposition answers this question under an additional assumption on $l(\cdot)$.

Proposition 4 *If $l(\cdot)$ is differentiable on $(0,\infty)$, then $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ is weakly increasing in b, either weakly decreasing or nonmonotone in c,[19] and strictly increasing in ϵ.*

The intuition for the comparative statics of $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ with respect to b or c depends on whether the C, $\underline{x}$-InOut-PEL, or $\overline{x}$-InOut-EL default is optimal. When the C default is optimal, a small change in one of these parameters does not affect the shaded area in Figure 4.

When the $\underline{x}$-InOut-PEL default is optimal, consider Figure 5. A small increase in b shifts up slightly Δ_L (see Figure 1) and, hence, $\bar{\Delta}^*(b,c,\epsilon) = \Delta_L + 2\epsilon$. This results in

[17] For each parameter, increasing it makes one of the inequalities in the second case in expression (3) easier to satisfy whilst making the other inequality in the second case in expression (3) harder to satisfy. As a result, it is not the case for any parameter that increasing it unambiguously contracts or expands the set of pairs of the other two parameters for which the $\underline{x}$-InOut-PEL default is optimal.

[18] The appendix considers additional relationships between the parameters and the optimal default.

[19] The proof shows that, at least when $l(z) = z^2$, (i) there exist (b,ϵ)-pairs for which $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ is weakly decreasing in c and (ii) there exist (b,ϵ)-pairs for which $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ is nonmonotone in c.

eliminating a sliver of shaded area just to the right of Δ_L and expanding the shaded area just to the right of $\Delta_L + 2\epsilon$ in Figure 5. Given that $L(\cdot)$ has higher values around $\Delta_L + 2\epsilon$ than around Δ_L, the net effect on the shaded area is positive, so that the total loss increases. Analogously, a small increase in c shifts down slightly Δ_L (see Figure 1) and, hence, $\bar{\Delta}^*(b, c, \epsilon) = \Delta_L + 2\epsilon$, so that the total loss decreases.

When the $\overline{x}$-InOut-EL default is optimal, consider Figure 6. A small increase in b has two effects. First, it shifts up slightly Δ_L, so that employees just to the right of Δ_L in the figure opt out. Because $L(\Delta_L) < c + l(b)$, these employees' losses increase. Second, it shifts the $c + l(b)$ line up, which increases the loss of employees with Δ's in $[\Delta_1 - 2\epsilon, \Delta_L]$, thus reinforcing the first effect.[20] A small increase in c has analogous effects, except that the first effect on the total loss is negative because Δ_L shifts slightly down rather than up. As a result the two effects have opposite signs and the overall effect of an increase in c on the total loss is ambiguous.

To see the intuition for why $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ is strictly increasing in ϵ, observe the following. First, $\frac{1}{2\epsilon}\int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta)d\Delta$ equals the average height of $L(\cdot)$ over the interval $[\bar{\Delta}^*(b, c, \epsilon) - 2\epsilon, \bar{\Delta}^*(b, c, \epsilon)]$. Second, in Figure 4/5/6 increasing ϵ expands the interval $[\bar{\Delta}^*(b, c, \epsilon) - 2\epsilon, \bar{\Delta}^*(b, c, \epsilon)]$ at both of its margins/at its right margin/at its left margin, where the value of $L(\cdot)$ exceeds the average value of this function over $[\bar{\Delta}^*(b, c, \epsilon) - 2\epsilon, \bar{\Delta}^*(b, c, \epsilon)]$.

The bottom line is that if the planner can lower employees' bias (e.g., through education) or decrease heterogeneity in the targeted population (e.g., by customising different defaults for, say, young vs. old employees), doing so is beneficial (or at least doesn't hurt). If the planner can affect the opt-out cost (e.g., by making employees jump through more/fewer hoops to opt out), she may want to either increase it or decrease it depending on which of the three cases in expression (3) applies and, if the

[20] The increase in b also raises Δ_1 (see Figure 3), which results in transferring a sliver of shaded area from just to the right of $\Delta_1 - 2\epsilon$ to just to the right of Δ_1 in Figure 6. Because $L(\Delta_1 - 2\epsilon) = L(\Delta_1)$, this effect is approximately 0.

third case applies, depending on which of the two effects mentioned above dominates.

6 Assumptions and Omitted Factors

In this section, we discuss some of the model's explicit and implicit assumptions. Many of the implicit assumptions are about potentially relevant factors that are omitted from the model.

6.1 Loss from Deviations and Distribution of Optimal Options

Propositions 2-4 assume that $l(\cdot)$ is strictly convex and symmetric. Proposition 4 assumes that $l(\cdot)$ is also differentiable on $(0, \infty)$. We also implicitly assume throughout that the loss from deviations depends on x and x' only through the difference $x - x'$. For brevity, let us call this the difference assumption.

Although such assumptions are common in the literature, they are not derived from deeper assumptions about utility functions, intertemporal budget constraints, and how employees might change in future periods the option with which they end up today. Notably, in the context of contribution rates, employer-match caps and tax-bracket thresholds are likely to lead to violations of symmetry, differentiability, and the difference assumption.[21]

Propositions 2-4 also assume a uniform distribution of optimal options, which is also clearly a simplification. Notably, this assumption excludes the realistic possibility

[21] For example, suppose there is an employer-match cap of m and let $\tilde{l}(x, x')$ denote the loss from deviations for an employee with optimal contribution rate x who ends up with contribution rate x'. In that case, $\tilde{l}(x, \cdot)$ probably exhibits a kink at m (so that differentiability fails) without a matching kink at $2x - m$ (so that symmetry around x fails). Moreover, given that the location of the kink does not depend on x, its position relative to x varies (so that the difference assumption fails). (I suspect that the conclusion of Proposition 4 continues to hold if $l(\cdot)$ is not differentiable at finitely many points, so that any failure of differentiability due to employer-match caps and tax-bracket thresholds may not be a problem.)

that there is a positive mass of employees with (i) optimal contribution rate equal to zero or to the employer-match cap or (ii) optimal allocation to stocks that equals zero or one.[22]

6.2 Normatively Irrelevant Opt-Out Costs

Within a standard model, employees' empirical reluctance to switch from default contribution rates implies opt-out costs in the thousands of dollars. Such opt-out costs seem orders of magnitude higher than what is justified based on the value of employees' time.[23]

Motivated by these observations, Goldin and Reck (2018) allow the planner to count a part of each employee's opt-out cost as normatively irrelevant and to exclude it from the social welfare function. Goldin and Reck also consider an extension of their baseline model, which allows employees to be biased. In that extension, the authors (i) derive (an approximation of) the first-order condition that an optimal default must satisfy, (ii) discuss whether marginally raising or lowering a default that is optimal in the absence of any bias improves welfare when a bias is present and (iii) discuss how the presence of a bias might undermine the case for AD.

The model in the current paper is less general than Goldin and Reck's extended model. Notably, our planner necessarily counts employees' full opt-out costs in the social welfare function. However, the lower generality makes it possible to perform a more complete analysis by explicitly solving for the optimal default policy and by exploring how the parameters affect the optimal default policy and the total loss in the population at the optimal default policy.

[22]The possibility of a positive mass of employees with optimal contribution rate equal to zero also strains the assumption that $b \leq \underline{x}$.

[23]See Goldin and Reck (2018) and Section IIB in Bernheim et al. (2015).

6.3 Updating from the Default

In the context of asset allocation, there is evidence that, upon seeing a default, employees update their preferred options in the direction of the default. In particular, Madrian and Shea (2001) find that employees hired before automatic enrolment for new hires allocated three times as much to new hires' default investment fund if they chose their asset allocation after the introduction of automatic enrolment for new hires. Madrian and Shea also report evidence that employees hired after automatic enrolment who had opted out of the default contribution rate or default investment fund were still much more heavily invested in the default investment fund than employees hired before automatic enrolment.

In the context of contribution rates, the evidence is more mixed. On the one hand, Madrian and Shea (2001) find that employees hired before automatic enrolment for new hires had contribution rates that were unaffected by whether they chose their contribution rates before or after the introduction of automatic enrolment for new hires. On the other hand, if employees sufficiently update their preferred contribution rates based on the default, this would explain their reluctance to opt out of the default contribution rate without any need to invoke exorbitant opt-out costs (see Section II D in Bernheim et al. (2015)).[24]

Ivanov (2019) modifies the model in the current paper by allowing for updating from the default under the assumption of a linear loss from deviations.[25] One finding in that paper is that, in addition to the three defaults here, two other defaults emerge as candidates for being the optimal default.[26] Another key insight is about how updating from the default can make AD the optimal default policy. Yet another

[24]Thus, in the presence of sufficient updating from the default, there is no need to consider a part of opt-out costs normatively irrelevant.

[25]In a second extension to their baseline model, Goldin and Reck (2018) also allow for updating from the default (but do not allow for any bias).

[26]The $\overline{x}$-InOut-EL default here corresponds to the $\underline{x}$-Out-$\overline{x}$-In-EL default in Ivanov (2019).

result is that, in contrast with Proposition 4, the total loss in the population at the optimal default policy can be nonmonotone in b and ϵ. A limitation of Ivanov (2019) is that the analysis is very complex and few simple intuitions for the main results emerge.

One advantage of the current paper relative to Ivanov (2019) is that the analysis here is more straightforward. Another advantage is that we can assume a strictly convex loss from deviations, which is probably more realistic than a linear loss from deviations. Of course, these advantages come at the cost of assuming away updating from the default. However, given the evidence cited above, such an assumption may be realistic in the context of contribution rates.

6.4 Procrastination

Adopting $\beta = 1$ as the normative benchmark, Carrol et al. (2009) solve for the optimal default policy in a model in which employees with β-δ preferences procrastinate opting out of suboptimal defaults. Given that the model in the current paper is static, it cannot address the issue of procrastination. On the other hand, in Carroll et al. (2009), employees are not biased regarding their optimal options.

6.5 Normative Ambiguity

The model here assumes that, for each employee, there is an option x, which is unambiguously optimal for her. This assumption is nontrivial. For example, suppose employees are dynamically inconsistent, so that they save differently (i) if they decide in each period how much to save for that period vs. (ii) if they can commit in the current period how much to save in the current period and in all future periods. Although many papers assume that the savings decisions under (i) are biased and

under (ii) are correct, the justification for this is unclear.[27]

Bernheim and Rangel (2009) develop a framework for conducting welfare analysis in the case of such normative ambiguity. Bernheim et al. (2015) apply this framework to the issue of optimal default policies in the context of contribution rates. Bernheim et al. allow for normative ambiguity regarding whether employees should opt out of the default. However, their analysis doesn't allow us to say much about the optimal default policy if employees misperceive their optimal options or if there is normative ambiguity regarding their optimal options.[28]

6.6 Decision Costs

We interpreted c as reflecting only implementation costs. Thus, we were implicitly assuming that employees face no decision costs, i.e., the time and cognitive costs of collecting information and thinking about what is the optimal option. Given the plausibility of nontrivial decision costs,[29] it is tempting to reinterpret c as reflecting both implementation and decision costs. However, this is problematic because of the following features of the model. First, all employees staying with a default avoid incurring c. However, this rules out the possibility that some employees incur the decision costs and conclude that staying with the default is a good idea. Second, employees' preferred options are not affected by whether they incurred the decision costs, which seems unrealistic.

There are papers which explore the implications of decision costs for the optimal

[27]See Bernheim and Rangel (2009) and Bernheim (2009).

[28]In their model with partially naive β-δ employees and their model with inattentive employees, each employee has a unique, unbiased preferred contribution rate. In their anchoring model, unless we are willing to assume that each employee correctly perceives her optimal option under AD but misperceives her optimal option when faced with a default, there is normative ambiguity regarding each employee's preferred contribution rate. However, this normative ambiguity prevents us from saying much about the optimal default policy.

[29]Based on a field experiment in Afghanistan, Blumenstock et al. (2018) conclude that decision costs are a major reason for why the default savings rate has a large effect on saving in a savings account.

default policy. In Carlin et al. (2013) and Wisson (2016), each individual doesn't know her optimal option, but has an unbiased belief about it and can learn about it if she incurs a cost. In Carlin et al. (2013), defaults are informative and, thus, discourage individuals from learning about their optimal options.[30] Because learning involves a positive externality, AD may be optimal. In Wisson (2016), individuals with β-δ preferences are excessively reluctant to incur the cost of learning or free-ride by letting others incur this cost. As a result, the optimal default policy may be an extreme default that induces individuals to incur the cost of learning.

6.7 The Presence of Sophisticated Employees

In the model here, all employees share the same parameters b and c. Such homogeneity is probably unrealistic.

One way to incorporate heterogeneity is to assume that a fraction, ρ, of the population is sophisticated in the sense that they have zero bias. Sophisticated employees can also have their own opt-out cost parameter, which would be smaller than the opt-out cost parameter for unsophisticated employees. Given that sophisticated employees would be less affected by the default in such a model, a natural guess is that the optimal default policy would be geared more towards the interests of unsophisticated employees, at least if ρ is not too high.

7 Concluding Remarks

The current paper analyses a model in which employees are biased regarding their optimal options. There are four main findings. First, AD is never an optimal default policy. This is largely due to the downward jump of the $L(\cdot)$ function at Δ_L. Second, given the values of the parameters, the optimal default is one of three defaults: the

[30] A similar mechanism is at work in Caplin and Martin (2017).

default placed in the middle of $[\underline{x}, \overline{x}]$ (the C default); the default, such that employees with $x = \underline{x}$ think it's too high, but are only just willing to stay with it (the $\underline{x}$-InOut-PEL default); and the default below $\overline{x}$, such that employees with $x = \overline{x}$ incur a loss from staying equal to the loss they would incur from opting out (the $\overline{x}$-InOut-EL default). The downward jumps of the $L(\cdot)$ function at Δ_L and Δ_R play a key role in determining the optimal default. Third, loosely speaking, increasing b or ϵ tilts the planner away from the C default and towards the $\overline{x}$-InOut-EL default, whereas increasing c has the opposite effect. Fourth, the total loss in the population at the optimal default is weakly increasing in b, either weakly decreasing or nonmonotone in c, and strictly increasing in ϵ.

Let us conclude with some cautionary remarks. First, as section 6 makes clear, the model in the current paper lacks realism in many ways, so that any conclusions should be taken with caution.

Second, any concrete policy advice would require estimating the model parameters as well as any parameters in a richer model from which $l(\cdot)$ can be derived. Deciding what value of b to use is likely to be especially hard. For example, there is no consensus in the existing literature on whether people undersave (see footnote 4). The question of whether people underweight stocks in their pension plan portfolios also hinges on complex issues, such as whether there is an equity premium puzzle, what is a reasonable expected return on the stock market going forward, and individuals' levels of risk aversion.

References

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2016. "Rich pickings? Risk, return, and skill in the portfolios of the wealthy." working paper.

Benartzi, Shlomo, and Richard H. Thaler. 1999. "Risk aversion or myopia? Choices in repeated gambles and retirement investments." *Management Science* 45.3: 364-381.

Bernheim, B. Douglas. 2009. "Behavioral Welfare Economics." *Journal of the European Economic Association* 7 (2-3): 267319.

Bernheim, B. Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics* 124 (1): 51104.

Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov. 2015. "The welfare economics of default options in 401 (k) plans." *American Economic Review* 105.9: 2798-2837.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2008. "The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States." In: Kay SJ, Sinha T Lessons from Pension Reform in the Americas. Oxford University Press; p. 59-87.

Beshears J, J.Choi J, Laibson D, Madrian BC. 2017. "Does Aggregated Returns Disclosure Increase Portfolio Risk-Taking?" *Review of Financial Studies* 30(6):1971-2005.

Blumenstock, JE, Callen, M, and Ghani, T. 2018. "Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan." *American Economic Review*, 108

(10), 2868-2901

Bogachev, Vladimir I. "Measure theory. Vol. 1." Springer Science & Business Media, 2007.

Bronchetti, Erin Todd, Thomas S. Dee, David B. Huffman, and Ellen Magenheim. 2013. "When a Nudge Isnt Enough: Defaults and Saving among Low-Income Tax Filers." *National Tax Journal* 66 (3): 60934.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. 2007. "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes." *Journal of Political Economy* 115.5: 707-747.

Campbell John Y. 2006. "Household Finance, Presidential Address to the American Finance Association." *Journal of Finance* LXI (4):1553-1604.

Caplin, Andrew, and Daniel Martin. 2017. "Defaults and Attention: The Drop Out Effect." *Revue conomique* 5: 45-54.

Carlin, Bruce Ian, Simon Gervais, and Gustavo Manso. 2013. "Libertarian paternalism, information production, and financial decision making." *The Review of Financial Studies* 26.9: 2204-2228.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 163974.

Chetty, Raj, John N. Friedman, Sren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings." *Quarterly Journal of Economics* 129.3: 1141-1219

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2002. "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance." In *Tax Policy and the Economy*, edited by James M. Poterba, 67113. Cambridge, MA: MIT Press.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2003. "Passive Decisions and Potent Defaults. National Bureau of Economic Research Working Paper 9917.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2004a. "For Better or Worse: Default Effects and 401(k) Savings Behavior." *In Perspectives on the Economics of Aging*, edited by David A. Wise, 81126. Chicago: University of Chicago Press.

Choi, James, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2004b. "Employees Investment Decisions About Company Stock." In *Pension Design and Structure: New Lessons from Behavioral Finance*, edited by Olivia Mitchell and Stephen Utkus. Oxford: Oxford University Press.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2006. "Saving for Retirement on the Path of Least Resistance." In *Behavioral Public Fi-*

nance: Toward a New Agenda, edited by Edward J. McCaffery and Joel Slemrod, 30451. New York: Russell Sage Foundation.

Eisenstein, Eric M., and Stephen J. Hoch. 2005. "Intuitive compounding: framing, temporal perspective, and expertise." Manuscript, Temple University.

Ericson KM, Laibson D. 2019. "Intertemporal Choice." In: Bernheim D, Laibson D, DellaVigna S, "Handbook of Behavioral Economics - Foundations and Applications 2." Elsevier

Goldin, Jacob and Daniel Reck. 2018. "Optimal Defaults with Normative Ambiguity." Working paper.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa. 2011. "IQ and stock market participation." *The Journal of Finance* 66.6: 2121-2164.

Ivanov, Asen. 2019. "Optimal Pension Plan Default Policies when Employees are Biased and Update from the Default." Working paper.

Levy, Matthew, and Joshua Tasoff. 2016. "Exponential-Growth Bias and Lifecycle Consumption." *Journal of the European Economic Association* 14.3

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 114987.

Poterba, James M. 2015. "Saver heterogeneity and the challenge of assessing retirement saving adequacy." *National Tax Journal* 68.2: 377.

Stango, Victor, and Jonathan Zinman. 2009. "Exponential Growth Bias and Household Finance." The Journal of Finance 64.6: 2807-2849.

Thaler, Richard H, and Cass R Sunstein. 2003. "Libertarian Paternalism." American Economic Review Papers and Proceedings 93 (2): 175-179.

Wisson, James. 2016. "Optimal Defaults and Uncertainty." Working paper.

8 Appendix: How Do the Parameters Affect the Optimal Default? Further Analysis.

The conditions in expression (3) are written so as to make clear the relative positions of various points on the horizontal axes in Figures 4-6. However, for the purposes of the current section, it is more convenient to rewrite expression (3) as follows:

$$\bar{\Delta}^*(b,c,\epsilon) = \begin{cases} \epsilon & \text{if } \epsilon \leq -\Delta_L \\ \Delta_L + 2\epsilon & \text{if } -\Delta_L < \epsilon \leq \frac{\Delta_1 - \Delta_L}{2} \\ \Delta_1 & \text{if } \epsilon > \frac{\Delta_1 - \Delta_L}{2} \end{cases} . \tag{4}$$

Letting $l^{-1}(\cdot)$ denote the inverse of $l(\cdot)$ on $(0,\infty)$, we have $\Delta_L = -l^{-1}(c) + b$ (see Figure 1). The following lemma, in conjunction with Lemma 2, will be useful in discussing how the parameters affect the conditions in expression (4).

Lemma 3 *Holding c fixed:*

1) $\lim_{b\downarrow 0} \frac{\Delta_1-\Delta_L}{2} = l^{-1}(c)$*;*

2) $\lim_{b\to\infty} \frac{\Delta_1-\Delta_L}{2} \geq \frac{l^{-1}(c)}{2}$.

Figure 7 shows the optimal default as a function of b and ϵ, holding c fixed. In particular, the C default is optimal below the line plotting $-\Delta_L$ as a function of b; the $\overline{x}$-InOut-EL default is optimal above the line plotting $\frac{\Delta_1-\Delta_L}{2}$ as a function of b; the $\underline{x}$-InOut-PEL default is optimal between the two lines. Based on this figure, the following relationships hold between the parameters and the optimal default.

1. Fixing c and ϵ, such that $\epsilon \leq \lim_{b\to\infty} \frac{\Delta_1-\Delta_L}{2}$, there is a cutoff value of b, such that the C/$\underline{x}$-InOut-PEL default is optimal for b below/above the cutoff.

2. Fixing c and ϵ, such that $\lim_{b\to\infty} \frac{\Delta_1-\Delta_L}{2} < \epsilon < l^{-1}(c)$, there are two cutoff values of b, such that the C/$\underline{x}$-InOut-PEL/$\overline{x}$-InOut-EL default is optimal for b below the lower cutoff/between the two cutoffs/above the higher cutoff.

3. Fixing c and ϵ, such that $\epsilon \geq l^{-1}(c)$, the $\overline{x}$-InOut-EL default is optimal for all b.

4. Fixing c and b, such that $b < l^{-1}(c)$, there are two cutoff values of ϵ, such that the C/$\underline{x}$-InOut-PEL/$\overline{x}$-InOut-EL default is optimal for ϵ below the lower cutoff/between the two cutoffs/above the higher cutoff.

5. Fixing c and b, such that $b \geq l^{-1}(c)$, there is a cutoff value of ϵ, such the $\underline{x}$-InOut-PEL/$\overline{x}$-InOut-EL default is optimal for ϵ below/above the cutoff.

6. Fixing b and ϵ, there are two cutoff values of c, such that the C/$\underline{x}$-InOut-PEL/$\overline{x}$-InOut-EL default is optimal for c above the higher cutoff/between the two cutoffs/below the lower cutoff. (To see this, observe that increasing c increases the $l^{-1}(c)$ intercept on each axis in Figure 7.)

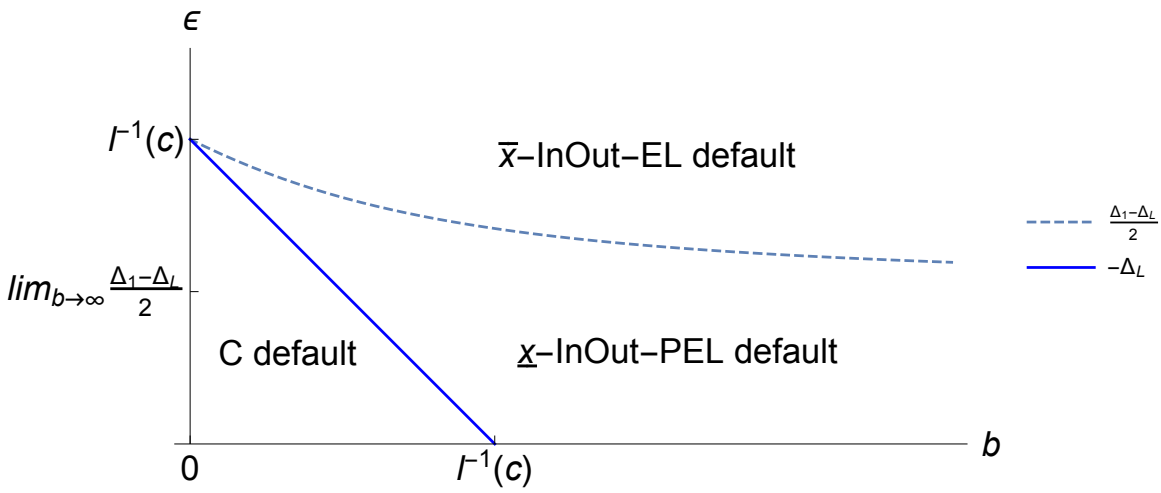

Figure 7: Optimal default as a function of b and ϵ, holding c fixed.

9 Appendix: Proofs

9.1 Proof of Lemma 1

The logic given straight after the lemma in the main text is rigorous. Nevertheless, it might be worth restating in more formal notation the arguments given there.

If $\Delta_L \leq 0$, we have $L(\Delta_L) = l(\Delta_L) < l(\Delta_L - b) = c$. The inequality follows because $\Delta_L - b < \Delta_L \leq 0$ and $l(\cdot)$ is strictly decreasing over $(-\infty, 0]$.

If $\Delta_L > 0$, we have $L(\Delta_L) = l(\Delta_L) < l(b)$. The inequality follows because $0 < \Delta_L < b$ and $l(\cdot)$ is strictly increasing over $[0, \infty)$.

Finally, $L(\Delta_R) = l(\Delta_R) = c + l(\Delta_R) - l(\Delta_R - b) > c + l(b) - l(0) = c + l(b)$. The second equality holds because $l(\Delta_R - b) = c$ (by the definition of Δ_R). The inequality holds because (i) the interval $[\Delta_R - b, \Delta_R]$ represents a rightward translation of the interval $[0, b]$ and (ii) by the strict convexity of $l(\cdot)$, the increase in $l(\cdot)$ must be larger over the former interval. Q.E.D.

9.2 Proof of Proposition 1

The logic given straight after the proposition is rigorous. Here, we just need to add that it implicitly assumes that there is a positive mass of employees in $[\overline{x}-\eta, \overline{x}]$. However, this assumption is without loss of generality–if it fails, just redefine $\overline{x}$ to be the upper bound of the support of the distribution of x's. Q.E.D.

9.3 Proof of Proposition 2

Let us first show that $\bar{\Delta}$, such that $\bar{\Delta} < \Delta_L$ or $\bar{\Delta} > \Delta_R$, cannot be a solution to problem (2) because setting $\bar{\Delta} = \Delta_1$ is strictly better. If $\bar{\Delta} < \Delta_L$ or $\bar{\Delta} > \Delta_R$, we have

$$\int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]} L(\Delta)d\Delta =$$
$$\int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta + \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c} L(\Delta)d\Delta =$$
$$\int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta + \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c} L(\Delta)d\Delta +$$
$$\int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta - \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta =$$
$$\int_{[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta + \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c} L(\Delta)d\Delta - \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta >$$
$$\int_{[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta + \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c} (c+l(b))d\Delta - \int_{[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]} (c+l(b))d\Delta =$$
$$\int_{[\Delta_1-2\epsilon,\Delta_1]} L(\Delta)d\Delta,$$

where the superscript "c" on an interval denotes its complement. The last equality

holds because because the two intervals $[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c$ and $[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]$ have equal length. To see that the inequality holds, let's consider separately the cases (i) $\bar{\Delta}<\Delta_L$ or $\bar{\Delta}-2\epsilon\geq\Delta_R$ and (ii) $\bar{\Delta}>\Delta_R$ and $\bar{\Delta}-2\epsilon<\Delta_R$. In case (i), $L(\Delta)=c+l(b)$ for all $\Delta\in[\bar{\Delta}-2\epsilon,\bar{\Delta}]$ (and, hence, for all $\Delta\in[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c$), $L(\Delta)\leq c+l(b)$ for all $[\Delta_1-2\epsilon,\Delta_1]$ (and, hence, for all $\Delta\in[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]$), and $L(\Delta)<c+l(b)$ for a positive measure of Δ's in $[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]$. In case (ii), $L(\Delta)\geq c+l(b)$ for all $\Delta\in[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c$, $L(\Delta)>c+l(b)$ for a positive measure of Δ's in $[\bar{\Delta}-2\epsilon,\bar{\Delta}]\cap[\Delta_1-2\epsilon,\Delta_1]^c$, and $L(\Delta)\leq c+l(b)$ for all $[\Delta_1-2\epsilon,\Delta_1]$ (and, hence, for all $\Delta\in[\bar{\Delta}-2\epsilon,\bar{\Delta}]^c\cap[\Delta_1-2\epsilon,\Delta_1]$). Thus, we can safely assume that any solution to problem (2) is in $[\Delta_L,\Delta_R]$.

Next, we show that $\bar{\Delta}^*(b,c,\epsilon)$ given in expression (3) is the unique solution to problem (2). We have:

$$\begin{aligned}
&\int_{\bar{\Delta}-2\epsilon}^{\bar{\Delta}} L(\Delta)d\Delta = \\
&\int_{-\infty}^{\bar{\Delta}} L(\Delta)d\Delta - \int_{-\infty}^{\bar{\Delta}-2\epsilon} L(\Delta)d\Delta = \\
&\int_{-\infty}^{\bar{\Delta}} L(\Delta)d\Delta - \int_{-\infty}^{\bar{\Delta}} L(\Delta-2\epsilon)d\Delta = \\
&\int_{-\infty}^{\bar{\Delta}} \big(L(\Delta)-L(\Delta-2\epsilon)\big)d\Delta.
\end{aligned}$$

It is apparent from Figures 4-6 that $L(\Delta)<L(\Delta-2\epsilon)$ for all $\Delta_L\leq\Delta<\bar{\Delta}^*(b,c,\epsilon)$. Thus, $\int_{-\infty}^{\bar{\Delta}}\big(L(\Delta)-L(\Delta-2\epsilon)\big)d\Delta$ is strictly decreasing in $\bar{\Delta}$ over $[\Delta_L,\bar{\Delta}^*(b,c,\epsilon)]$. Similarly, it is apparent from Figures 4-6 that $L(\Delta)>L(\Delta-2\epsilon)$ for all $\bar{\Delta}^*(b,c,\epsilon)<\Delta\leq\Delta_R$. Thus, $\int_{-\infty}^{\bar{\Delta}}\big(L(\Delta)-L(\Delta-2\epsilon)\big)d\Delta$ is strictly increasing in $\bar{\Delta}$ over $[\bar{\Delta}^*(b,c,\epsilon),\Delta_R]$. Thus, $\int_{\bar{\Delta}-2\epsilon}^{\bar{\Delta}} L(\Delta)d\Delta$ has a unique minimum over $[\Delta_L,\Delta_R]$ at $\bar{\Delta}=\bar{\Delta}^*(b,c,\epsilon)$.

It remains to verify that $\bar{\Delta}^*(b,c,\epsilon)$ satisfies $\overline{x}-1\leq\bar{\Delta}^*(b,c,\epsilon)\leq\overline{x}$. The first

inequality holds because $\bar{\Delta}^*(b, c, \epsilon) > 0$ (see Figures 4-6) and $\overline{x} \leq 1$. In the first case in expression (3), $\bar{\Delta}^*(b, c, \epsilon) = \epsilon < \overline{x}$. In the second case in expression (3), $\bar{\Delta}^*(b, c, \epsilon) = \Delta_L + 2\epsilon < b + 2\epsilon \leq \overline{x}$, where the last inequality follows from our assumption $b \leq \underline{x}$. In the third case in expression (3), $\bar{\Delta}^*(b, c, \epsilon) = \Delta_1 = \Delta_1 - \Delta_L + \Delta_L < 2\epsilon + \Delta_L < 2\epsilon + b \leq \overline{x}$, where the last inequality follows from our assumption $b \leq \underline{x}$. Q.E.D.

9.4 Proof of Lemma 2

The logic given straight after the lemma in the main text is rigorous. Nevertheless, it might be worth restating in more formal notation the arguments given there for why $\Delta_1 - \Delta_L$ is strictly decreasing in b.

Let us write $\Delta_L(b, c)$ and $\Delta_1(b, c)$ to make explicit the dependence of Δ_L and Δ_1 on b and c. Given the fact that $\Delta_1(b, c) > b$ and the convexity of $l(\cdot)$, for any $\sigma > 0$ we have:

$$l(\Delta_1(b, c) + \sigma) - l(\Delta_1(b, c)) > l(b + \sigma) - l(b),$$

so that

$$l(\Delta_1(b, c) + \sigma) > l(\Delta_1(b, c)) + l(b + \sigma) - l(b) = c + l(b + \sigma),$$

where the equality follows from $l(\Delta_1(b, c)) = c + l(b)$.

Given that $l(\Delta_1(b, c)) < c + l(b + \sigma) < l(\Delta_1(b, c) + \sigma)$, it must be that $\Delta_1(b, c) < \Delta_1(b + \sigma, c) < \Delta_1(b, c) + \sigma$. Thus,

$$\Delta_1(b + \sigma, c) - \Delta_L(b + \sigma, c) - (\Delta_1(b, c) - \Delta_L(b, c)) < \sigma - \Delta_L(b + \sigma, c) + \Delta_L(b, c) = 0.$$

The equality follows from the fact that increasing the bias from b to $b+\sigma$ increases Δ_L one-for-one (see Figure 1). Q.E.D.

9.5 Proof of Lemma 3

Statement 1) follows because, as b goes to 0, Δ_1 is arbitrarily close to the positive root of $l(\Delta) = c$ (see Figure 3) and $\Delta_L = -l^{-1}(c) + b$ is arbitrarily close to $-l^{-1}(c)$.

Statement 2) hold because $\Delta_1 - \Delta_L$ is strictly decreasing in b (see Lemma 2) and $\Delta_1 - \Delta_L \geq b - \Delta_L = l^{-1}(c)$. Q.E.D.

9.6 Proof of Proposition 3

Let us write $\Delta_L(b, c)$ and $\Delta_1(b, c)$ to make explicit the dependence of Δ_L and Δ_1 on b and c.

Given Lemma 2 and the fact that $\Delta_L(b, c)$ is strictly increasing in b and strictly decreasing in c (see Figure 1), the statements in the proposition follow if we replace "contracts" and "expands" with "weakly contracts" and "weakly expands", respectively. It remains to show that the statements continue to hold without the "weakly" qualifier.[31]

First, consider statement 1) and suppose that $b' > b$. If c is such that $\Delta_L(b, c) < 0$[32] and ϵ is such that $\Delta_L(b, c) < -\epsilon < \Delta_L(b', c)$, the C default is optimal given (b, c, ϵ), but not given (b', c, ϵ). If c is arbitrary and ϵ is such that $\Delta_1(b', c) - \Delta_L(b', c) < 2\epsilon < \Delta_1(b, c) - \Delta_L(b, c)$, the $\overline{x}$-InOut-EL default is optimal given (b', c, ϵ), but not given (b, c, ϵ). This establishes statement 1) without the "weakly" qualifier. The proof of statement 2) without the "weakly" qualifier is analogous.

Next, consider statement 3) and suppose that $\epsilon' > \epsilon$. Let b be arbitrary. If c is such that $-\epsilon' < \Delta_L(b, c) < -\epsilon$,[33] the C default is optimal given (b, c, ϵ), but not given

[31] The terms "weakly contracts" and "weakly expands" are used in the sense of weak set inclusion.

[32] We can always find a value of c that is large enough, so that $\Delta_L(b, c) < 0$ (see Figure 1).

[33] Given that $\Delta_L(b, \cdot)$ is continuous and its range is $(-\infty, b)$ (see Figure 1), we can find for any b, ϵ, and $\epsilon' > \epsilon$ a value of c, such that $-\epsilon' < \Delta_L(b, c) \leq -\epsilon$.

(b, c, ϵ'). If c is such that $2\epsilon < \Delta_1(b, c) - \Delta_L(b, c) < 2\epsilon'$,[34] the $\overline{x}$-InOut-EL default is optimal given (b, c, ϵ'), but not given (b, c, ϵ). This establishes statement 3) without the "weakly" qualifier. Q.E.D.

9.7 Proof of Proposition 4

The proof makes use of a sequence of claims. Before we launch into these, let us introduce some additional notation.

Let $\Psi(b, c, \epsilon) = \frac{1}{2\epsilon} \int_{\bar{\Delta}^*(b,c,\epsilon)-2\epsilon}^{\bar{\Delta}^*(b,c,\epsilon)} L(\Delta) d\Delta$. Let Ψ_p denote the partial derivative of Ψ with respect to the parameter $p \in \{b, c, \epsilon\}$. Let $-p$ denote the two parameters other than p and, with the usual abuse of notation, we will write $\bar{\Delta}^*(p, -p)$, $\Psi(p, -p)$, and $\Psi_p(p, -p)$. Let $p(-p, i)$ denote the set of values of p that satisfy, given fixed $-p$, the inequalities in the ith case in expressions (3) and (4). Finally, let $l'(\cdot)$ denote the derivative of $l(\cdot)$ and let $l^{-1}(\cdot)$ denote the inverse of $l(\cdot)$ on $(0, \infty)$.

Claim 1

$$\Psi(b, c, \epsilon) = \frac{1}{2\epsilon} \times \begin{cases} \int_{-\epsilon}^{\epsilon} l(\Delta) d\Delta & \textit{if } \epsilon \leq -\Delta_L \\ \int_{\Delta_L}^{\Delta_L + 2\epsilon} l(\Delta) d\Delta & \textit{if } -\Delta_L < \epsilon \leq \frac{\Delta_1 - \Delta_L}{2} \\ (\Delta_L - \Delta_1 + 2\epsilon)(c + l(b)) + \int_{\Delta_L}^{\Delta_1} l(\Delta) d\Delta & \textit{if } \epsilon > \frac{\Delta_1 - \Delta_L}{2} \end{cases} . \tag{5}$$

Proof:

The claim follows directly from Figures 4-6. Q.E.D.

Claim 2

1) One of the following holds.

[34] Given that $\Delta_1(b, \cdot) - \Delta_L(b, \cdot)$ is continuous and its range is $(0, \infty)$ (see Figure 3), we can find for any b, ϵ, and $\epsilon' > \epsilon$ a value of c, such that $2\epsilon \leq \Delta_1(b, c) - \Delta_L(b, c) < 2\epsilon'$.

(i) $b(-b,1) = b(-b,2) = \emptyset$ *and* $b(-b,3) = (0,\infty)$ *or*

(ii) $b(-b,1) = (0,b_1]$, $b(-b,2) = (b_1,\infty)$, *and* $b(-b,3) = \emptyset$, *where* $b_1 > 0$, *or*

(iii) $b(-b,1) = (0,b_1]$, $b(-b,2) = (b_1,b_2]$, *and* $b(-b,3) = (b_2,\infty)$, *where* $0 < b_1 < b_2$.

2) $c(-c,3) = (0,c_1)$, $c(-c,2) = [c_1,c_2)$, *and* $c(-c,1) = [c_2,\infty)$, *where* $0 < c_1 < c_2$.

3) One of the following holds.

(i) $\epsilon(-\epsilon,1) = \emptyset$, $\epsilon(-\epsilon,2) = (0,\epsilon_1]$, *and* $\epsilon(-\epsilon,3) = (\epsilon_1,\infty)$, *where* $\epsilon_1 > 0$, *or*

(ii) $\epsilon(-\epsilon,1) = (0,\epsilon_1]$, $\epsilon(-\epsilon,2) = (\epsilon_1,\epsilon_2]$, *and* $\epsilon(-\epsilon,3) = (\epsilon_2,\infty)$, *where* $0 < \epsilon_1 < \epsilon_2$.

Proof:

The claim follows from expression (4) and Figure 7, which plots the $-\Delta_L$ and $\frac{\Delta_1-\Delta_L}{2}$ curves as functions of b, holding c fixed. (The key features of the curves in the figure are based on Lemmas 2 and 3.) In particular, looking at the figure, it is straightforward that statements 1) and 3) hold. To see that statement 2) also holds, observe that increasing c increases the $l^{-1}(c)$ intercept on each axis in Figure 7.[35] Q.E.D.

Claim 3 *Assume that* $l(\cdot)$ *is differentiable on* $(0,\infty)$. *Then, the following hold.*

1) $l'(\cdot) > 0$ *on* $(0,\infty)$.

2) Δ_L *and* Δ_1 *are continuously differentiable in each parameter.*

3) $\frac{\partial \Delta_L}{\partial c} < 0$.

[35] Figure 7 doesn't make clear whether various intervals in Claim 2 should include their endpoints. However, this can easily be ascertained by consulting expression (4).

Proof:

Given that $l(\cdot)$ is increasing on $(0,\infty)$, we have $l'(\cdot) \geq 0$ on $(0,\infty)$. Given that $l(\cdot)$ is strictly convex, $l'(\cdot)$ is strictly increasing, so that it cannot equal zero anywhere on $(0,\infty)$.

Next, consider statement 2). Given that $l(\cdot)$ is differentiable on $(0,\infty)$ and convex, it is continuously differentiable on $(0,\infty)$. Given that $l(\cdot)$ is continuously differentiable on $(0,\infty)$ and $l'(\cdot) \neq 0$ on $(0,\infty)$, $l^{-1}(\cdot)$ is continuously differentiable on $(0,\infty)$. Given that $\Delta_L = -l^{-1}(c) + b$ and $\Delta_1 = l^{-1}(c + l(b))$, Δ_L and Δ_1 are continuously differentiable in each parameter.

Statement 3) follows because $\Delta_L = -l^{-1}(c) + b$ and the derivative of $l^{-1}(\cdot)$ evaluated at c equals $1/l'(l^{-1}(c)) > 0$. Q.E.D.

Claim 4 *The maximand in problem (2) is continuous in* $(\bar{\Delta}, p)$.

Proof:

When $p = \epsilon$, the integrand does not depend on p. The claim follows from the continuity of the definite integral with respect to the bounds of integration.

From here on, we restrict attention to $p \in \{b, c\}$. We write $L(\Delta, p)$, $\Delta_L(p)$, and $\Delta_R(p)$ to make explicit the dependence on p of expression (1), Δ_L, and Δ_R, respectively.

Take $\sigma > 0$. We need to find $\delta > 0$, such that $|\int_{\bar{\Delta}-2\epsilon}^{\bar{\Delta}} L(\Delta, p)d\Delta - \int_{\bar{\Delta}'-2\epsilon}^{\bar{\Delta}'} L(\Delta, p')d\Delta| < \sigma$ whenever $(\bar{\Delta}', p')$ is within δ Euclidean distance of $(\bar{\Delta}, p)$.

Let $M = 1 + \sup_{\Delta} L(\Delta, p)$ and $r = min(\frac{\sigma}{8M}, 1)$. Also, let I, J and K denote the sets $[\bar{\Delta} - 2\epsilon, \bar{\Delta}]$, $[\bar{\Delta}' - 2\epsilon, \bar{\Delta}']$, and $[\Delta_L(p) - r/4, \Delta_L(p) + r/4] \cup [\Delta_R(p) - r/4, \Delta_R(p) + r/4]$, respectively. We have:

$$|\int_I L(\Delta, p)d\Delta - \int_J L(\Delta, p')d\Delta| =$$

$$|\int\limits_{I\setminus K} L(\Delta,p)d\Delta + \int\limits_{I\cap K} L(\Delta,p)d\Delta - \int\limits_{J\setminus K} L(\Delta,p')d\Delta - \int\limits_{J\cap K} L(\Delta,p')d\Delta| \leq$$
$$|\int\limits_{I\setminus K} L(\Delta,p)d\Delta - \int\limits_{J\setminus K} L(\Delta,p')d\Delta| + \int\limits_{I\cap K} L(\Delta,p)d\Delta + \int\limits_{J\cap K} L(\Delta,p')d\Delta \leq$$
$$|\int\limits_{I\setminus K} L(\Delta,p)d\Delta - \int\limits_{J\setminus K} L(\Delta,p')d\Delta| + \int\limits_{K} L(\Delta,p)d\Delta + \int\limits_{K} L(\Delta,p')d\Delta$$

Clearly, $\sup\limits_{\Delta\in K} L(\Delta,p) < M$. Also, we can choose δ small enough that $\sup\limits_{\Delta\in K} L(\Delta,p') < M$. In that case, the last expression is smaller than:

$$|\int\limits_{I\setminus K} L(\Delta,p)d\Delta - \int\limits_{J\setminus K} L(\Delta,p')d\Delta| + 2rM \leq$$
$$|\int\limits_{I\setminus K} L(\Delta,p)d\Delta - \int\limits_{J\setminus K} L(\Delta,p')d\Delta| + \frac{\sigma}{4} =$$
$$|\int\limits_{I\setminus(K\cup J)} L(\Delta,p)d\Delta + \int\limits_{(I\setminus K)\cap J} L(\Delta,p)d\Delta - \int\limits_{J\setminus(K\cup I)} L(\Delta,p')d\Delta - \int\limits_{(J\setminus K)\cap I} L(\Delta,p')d\Delta| + \frac{\sigma}{4} \leq$$
$$|\int\limits_{(I\setminus K)\cap J} L(\Delta,p)d\Delta - \int\limits_{(J\setminus K)\cap I} L(\Delta,p')d\Delta| + \int\limits_{I\setminus(K\cup J)} L(\Delta,p)d\Delta + \int\limits_{J\setminus(K\cup I)} L(\Delta,p')d\Delta + \frac{\sigma}{4} =$$
$$|\int\limits_{(I\cap J)\setminus K} (L(\Delta,p) - L(\Delta,p'))d\Delta| + \int\limits_{I\setminus(K\cup J)} L(\Delta,p)d\Delta + \int\limits_{J\setminus(K\cup I)} L(\Delta,p')d\Delta + \frac{\sigma}{4} \leq$$
$$|\int\limits_{(I\cap J)\setminus K} (L(\Delta,p) - L(\Delta,p'))d\Delta| + \int\limits_{I\setminus J} L(\Delta,p)d\Delta + \int\limits_{J\setminus I} L(\Delta,p')d\Delta + \frac{\sigma}{4}.$$

We can choose δ small enough, so that (i) $L(\Delta,p) - L(\Delta,p') < \frac{\sigma}{8\epsilon}$ on $(I\cap J)\setminus K$, (ii) the length of each of $I\setminus J$ and $J\setminus I$ is less than $\frac{\sigma}{4M}$, and (iii) $\sup\limits_{\Delta\in J\setminus I} L(\Delta,p') < M$. In that case, the last expression is less than

$$\frac{\sigma}{8\epsilon}2\epsilon + M\frac{\sigma}{4M} + M\frac{\sigma}{4M} + \frac{\sigma}{4} = \sigma.$$

Q.E.D.

Claim 5 *$\Psi(\cdot, -p)$ is continuous.*

Proof:

By Claim 4, the maximand in problem (2) is continuous in $(\bar{\Delta}, p)$. The claim follows from the Maximum Theorem applied to the special case when the optimisation problem has a unique solution and the constraints do not depend on the parameters. (Proposition 2 shows that problem (2) has a unique solution.) Q.E.D.

Claim 6 *Assume that (i) $f : (0, \infty) \to \mathbb{R}$ is absolutely continuous and weakly monotone on $[a, b]$, and (ii) $\phi : \mathbb{R} \to \mathbb{R}$ is absolutely continuous on $f([a, b])$.*[36] *Then, $\phi \circ f$ is absolutely continuous on $[a, b]$.*

Proof: Assume that f is weakly increasing on $[a, b]$.[37] Given $\epsilon > 0$, let $\delta > 0$ be such that, for any finite collection of pairwise disjoint intervals $(c_k, d_k) \subseteq f([a, b])$ satisfying $\sum_k (d_k - c_k) < \delta$, we have $\sum_k |\phi(d_k) - \phi(c_k)| < \epsilon$. Given this δ, let $\tau > 0$ be such that, for any finite collection of pairwise disjoint intervals $(a_k, b_k) \subseteq [a, b]$ satisfying $\sum_k (b_k - a_k) < \tau$, we have $\sum_k |f(b_k) - f(a_k)| < \delta$. We have $\sum_k |\phi(f(b_k)) - \phi(f(a_k))| = \sum_{\{k | f(a_k) \neq f(b_k)\}} |\phi(f(b_k)) - \phi(f(a_k))|$. Given that f is weakly increasing, any two intervals $(f(a_{k'}), f(b_{k'}))$ and $(f(a_{k''}), f(b_{k''}))$, such that $f(a_{k'}) \neq f(b_{k'})$ and $f(a_{k''}) \neq f(b_{k''})$, are disjoint. Thus, the latter sum is less than ϵ.[38] Q.E.D.

Claim 7 *Assume that $l(\cdot)$ is differentiable on $(0, \infty)$. Fix an arbitrary interval $[p', p'']$, where $0 < p' < p''$. $\Psi(\cdot, -p)$ is absolutely continuous on $[p', p'']$.*

Proof:

[36] $f([a, b])$ denotes the image of $[a, b]$ under f.
[37] The proof is analogous if f is weakly decreasing.
[38] This proof is based on the hint given at the end of problem 5.8.59 in Bogachev (2007).

Given that Δ_L and Δ_1 are continuously differentiable in p (see Claim 3), they are absolutely continuous on $[p', p'']$. Given that they are also weakly monotone in p and the definite integral is absolutely continuous in each bound of integration, it follows from Claim 6 that each integral in expression (5) is absolutely continuous in p on $[p', p'']$. It is straightforward that each piece in expression (5) is absolutely continuous on $[p', p'']$.[39]

The absolute continuity of $\Psi(\cdot, -p)$ on $[p', p'']$ follows from (i) the continuity of $\Psi(\cdot, -p)$ (see Claim 5), (ii) the absolute continuity of each piece in expression (5) on $[p', p'']$, and (iii) the fact that there are finitely many points in $[p', p'']$ at which there is a switch between the cases in expression (5) (see Claim 2). Q.E.D.

Claim 8 *Assume that $l(\cdot)$ is differentiable. Then, the following hold.*

1) For almost all b, $\Psi_b(b, c, \epsilon)$ exists and

$$\Psi_b(b, c, \epsilon) = \begin{cases} 0 & \text{if } b \in b(-b, 1) \\ \frac{1}{2\epsilon}(l(\Delta_L + 2\epsilon) - l(\Delta_L)) & \text{if } b \in b(-b, 2) \\ \frac{1}{2\epsilon}\left((\Delta_L - \Delta_1 + 2\epsilon)l'(b) + (c + l(b) - l(\Delta_L))\right) & \text{if } b \in b(-b, 3) \end{cases} . \tag{6}$$

2) For almost all c, $\Psi_c(b, c, \epsilon)$ exists and

$$\Psi_c(b, c, \epsilon) = \begin{cases} 0 & \text{if } c \in c(-c, 1) \\ \frac{1}{2\epsilon}(l(\Delta_L + 2\epsilon) - l(\Delta_L))\frac{\partial \Delta_L}{\partial c} & \text{if } c \in c(-c, 2) \\ \frac{1}{2\epsilon}\left((\Delta_L - \Delta_1 + 2\epsilon) + (c + l(b) - l(\Delta_L))\frac{\partial \Delta_L}{\partial c}\right) & \text{if } c \in c(-c, 3) \end{cases} . \tag{7}$$

[39] Given that $l(\cdot)$ is differentiable on $(0, \infty)$ and convex, it is continuously differentiable on $(0, \infty)$. Hence, the term $l(b)$ in the third piece in expression (5) is absolutely continuous in b.

3) For almost all ϵ, $\Psi_\epsilon(b,c,\epsilon)$ exists and

$$\Psi_\epsilon(b,c,\epsilon) = \frac{1}{\epsilon}\left(l(\bar{\Delta}^*(b,c,\epsilon)) - \Psi(b,c,\epsilon)\right). \tag{8}$$

Proof:

The claim follows by differentiating each piece in expression (5) with respect to p. The differentiation of any integrals is done via the Leibnitz integral rule. Given that Δ_L and Δ_1 are continuously differentiable in each parameter (see Claim 3), we can use this rule when the limits of integration involve Δ_L and Δ_1.[40] Q.E.D.

Proof of Proposition 4:

Suppose that p' and p'' are such that $0 < p' < p''$. By the absolute continuity of $\Psi(\cdot, -p)$ (see Claim 7), we can write:

$$\Psi(p'', -p) - \Psi(p', -p) = \int_{p'}^{p''} \Psi_p(\tilde{p}, -p)d\tilde{p}.$$

First, suppose that $p = b$. It is easy to see that each piece in expression (6) is nonnegative. Thus, for any nondegenerate $[b', b''] \subseteq (0, \infty)$, $\Psi(b'', c, \epsilon) - \Psi(b', c, \epsilon) = \int_{b'}^{b''} \Psi_b(\tilde{b}, c, \epsilon)d\tilde{b} \geq 0$.

Next, suppose that $p = c$. Whenever we wish to make the dependence of Δ_L and Δ_1 on c explicit, we will write $\Delta_1(c)$ and $\Delta_L(c)$. Let c_1 and c_2 be as in statement 2) in Claim 2.

For all $c \in [c_2, \infty)$, the first piece of expression (7) applies. Thus, for any nondegenerate $[c', c''] \subseteq [c_2, \infty)$, $\Psi(b, c'', \epsilon) - \Psi(b, c', \epsilon) = \int_{c'}^{c''} \Psi_c(b, \tilde{c}, \epsilon)d\tilde{c} = 0$.

[40] At points p_1 and p_2 in each statement in Claim 2, different pieces of expression (5) apply to the left and right, so that $\Psi_p(p, -p)$ may not exist at $p = p_1$ or $p = p_2$. This does not invalidate Claim 8 given the "for almost all" qualification.

For all $c \in [c_1, c_2)$, the second piece of expression (7) applies. This second piece is negative given that (i) $l(\Delta_L + 2\epsilon) - l(\Delta_L) > 0$ whenever $-\Delta_L < \epsilon$ and (ii) $\frac{\partial \Delta_L}{\partial c} < 0$ (see Claim 3). Thus, for any nondegenerate $[c', c''] \subseteq [c_1, c_2)$, $\Psi(b, c'', \epsilon) - \Psi(b, c', \epsilon) = \int_{c'}^{c''} \Psi_c(b, \tilde{c}, \epsilon) d\tilde{c} < 0$.

We have shown that $\Psi(b, \cdot, \epsilon)$ is weakly decreasing on $[c_1, \infty)$. It follows that $\Psi(b, \cdot, \epsilon)$ is weakly decreasing or nonmonotone on $(0, \infty)$, which is the statement in Proposition 4. However, here we go further by establishing that, depending on the values of b and ϵ, both possibilities can in fact occur if $l(z) = z^2$.[41]

Assuming that $l(z) = z^2$, we consider the behaviour of $\Psi(b, \cdot, \epsilon)$ on $(0, c_1)$. For all $c \in (0, c_1)$, the third piece of expression (7) applies. With a quadratic loss from deviations, this third piece becomes:

$$\frac{1}{2\epsilon}(\Delta_L(c) - \Delta_1(c) + 2\epsilon - b).$$

If $b \geq 2\epsilon$, we have $\Delta_L(c) - \Delta_1(c) + 2\epsilon - b < 0$. Thus, for any nondegenerate $[c', c''] \subseteq (0, c_1)$, $\Psi(b, c'', \epsilon) - \Psi(b, c', \epsilon) = \int_{c'}^{c''} \Psi_c(b, \tilde{c}, \epsilon) d\tilde{c} < 0$.

If $b < 2\epsilon$, we have $\lim_{c \downarrow 0}(\Delta_L(c) - \Delta_1(c) + 2\epsilon - b) = 2\epsilon - b > 0$. The equality follows because $\lim_{c \downarrow 0}(\Delta_L(c) - \Delta_1(c)) = 0$ (see Figure 3). Thus, there exists c_0, such that $0 < c_0 < c_1$ and $\Delta_L(c) - \Delta_1(c) + 2\epsilon - b > 0$ for all $c < c_0$. Thus, for any nondegenerate $[c', c''] \subseteq (0, c_0)$, $\Psi(b, c'', \epsilon) - \Psi(b, c', \epsilon) = \int_{c'}^{c''} \Psi_c(b, \tilde{c}, \epsilon) d\tilde{c} > 0$.

Finally, let us turn to the case $p = \epsilon$. Observe that (i) $\Psi(b, c, \epsilon)$ equals the average height of $L(\cdot)$ over the interval $[\bar{\Delta}^*(b, c, \epsilon) - 2\epsilon, \bar{\Delta}^*(b, c, \epsilon)]$ and (ii) $l(\bar{\Delta}^*(b, c, \epsilon))$ is strictly larger than this average height (see Figures 4-6). Hence, expression (8) is strictly positive. Thus, for any nondegenerate $[\epsilon', \epsilon''] \subseteq (0, \infty)$, $\Psi(b, c, \epsilon'') - \Psi(b, c, \epsilon') = \int_{\epsilon'}^{\epsilon''} \Psi_\epsilon(b, c, \tilde{\epsilon}) d\tilde{\epsilon} > 0$. Q.E.D.

[41] I suspect this is true for any $l(\cdot)$ satisfying the assumptions behind Proposition 4, but I have not proved this.

School of Economics and Finance

Queen Mary
University of London

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/