Default Options and Retirement Saving Dynamics

Taha Choukhmane[1]

June, 2021

Click here for the latest version

Abstract

I document that employees offset the short-run positive effect of auto-enrollment in retirement plans by saving less in the future. Consequently, a structurally-estimated lifecycle model predicts that the long-term effect of auto-enrollment on wealth is negligible except at the bottom of the earnings distribution. The observed inertia at the savings default is explained by an opt-out cost of around $250. My estimate is smaller than the thousands of dollars estimated in previous studies because non-autoenrolled workers can compensate for not contributing now by contributing more later. Auto-enrollment improves welfare if the policymaker is paternalistic or has strong redistributive preferences.

[1]Choukhmane: MIT Sloan School of Management (tahac@mit.edu). I am deeply indebted to my dissertation committee for their generous help at all stages of this project: James Choi, Nick Barberis, Costas Meghir, and Cormac O'Dea. For helpful discussions and suggestions, I thank Jason Abaluck, Joseph Altonji, Antoine Arnoud, Doug Bernheim, Nicolas Coeurdacier, Stefano DellaVigna, Eric French, Xavier Gabaix, John Eric Humphries, David Laibson, Patrick Moran, Yusuke Narita, Matthew Notowidigdo, Michael Peters, Daniel Reck, Jesse Shapiro, Jonathan Skinner, Jaya Wen, Nicolas Werquin, and Fabrizio Zilibotti. I thank an anonymous U.S. financial institution for making available the data used in this paper, and for helpful discussions and technical support. I acknowledge financial support from the Whitebox Advisors Fellowship, the Boston College Center for Retirement Research, and the Sloan Foundation.

1 Introduction

Automatic enrollment has transformed the way millions of people around the world save for retirement. Starting with Madrian and Shea (2001), several studies have found that moving from an opt-in regime (where new hires must actively elect to participate) to an automatic enrollment regime (where new hires are enrolled by default with the option to opt out) can double participation in a retirement savings plan.[2] Motivated by this evidence, auto-enrollment has become perhaps the most common policy application of behavioral economics and is already the norm in retirement savings plans in the U.S., U.K., New Zealand, and Turkey.

Does auto-enrollment increase lifetime wealth accumulation and welfare? Despite its widespread adoption, the answer to this question remains unknown. Auto-enrollment policies are too recent to directly observe the effect that enrolling a 25-year-old has on her retirement wealth 40 years later. In this paper, I show that considering automatic enrollment in a fully-specified dynamic setting is crucial for understanding its long-term savings and welfare impact. I document that employees offset the short-run positive effect of automatic enrollment by saving less in the future. Consequently, a lifecycle model predicts that the long-term effect of auto-enrollment on wealth is negligible except at the bottom of the lifetime earnings distribution. The estimated model fits the data well (both in and out of sample), and explains the observed inertia at the savings default with an opt-out cost of only around \$250 and otherwise standard preferences.

I begin by documenting three stylized facts about auto-enrollment which provide motivation for the model. First, using a new proprietary dataset of 401(k) administrative records from 86 U.S. firms, I document that changing a positive default contribution rate to an even higher default lowers the 401(k) participation rate. A model with a cost of opting out of the savings default fits this new evidence, while alternative theories, such as loss aversion and psychological anchoring, make the opposite prediction. This cost of opting out of the default could reflect both real transaction costs and a broad class of behavioral explanations of the default effect, such as the cognitive cost of paying attention or finding the optimal choice.

Second, I show that the median cumulative contributions of non-autoenrolled workers catch up to the median cumulative contributions of autoenrolled workers over three years. Because non-autoenrolled workers can compensate for their low savings early on by saving more later, they are

[2]Large auto-enrollment effects on participation and contributions have been documented in the U.S. (Choi et al., 2004 and 2006; Beshears et al., 2009; Clark et al., 2015), in the U.K. (Cribb and Emmerson, 2016), and in an experimental setting in Afghanistan (Blumenstock et al., 2018).

not fully forgoing the tax and matching benefits of retirement savings when they do not save early in their tenure. Hence, a modest opt-out cost can explain the observed level of inertia around the savings default.

Third, using representative data from the U.K., I find no evidence that auto-enrollment creates long-lasting saving habits and, if anything, previously auto-enrolled workers are less likely to participate when required to make an active opt-in decision in their next job. Specifically, I find that auto-enrolling workers in their current job's retirement savings plan caused their participation rate and contribution rate in their next employer's opt-in savings plan to fall by 13 percentage points and 0.5% of income, respectively. When the new employer has auto-enrollment, I find no difference in the contributions of previously auto-enrolled and not previously auto-enrolled new hires.

These stylized facts motivate a lifecycle model with opt-out costs. I develop a structural model in which individuals save in both a realistic retirement account and a liquid asset, borrow through unsecured credit, and face labor market risk and a progressive tax and benefit system. I exploit quasi-experimental variation in the default contribution rate at 34 U.S. 401(k) plans to identify the model's preference parameters. Specifically, I compare the contribution decisions of employees hired in the year before and the year after the introduction of auto-enrollment for new hires.

I estimate an opt-out cost of around $250 every time an employee changes her contribution rate. Previous work, which has not studied auto-enrollment in a fully-specified dynamic model, infers opt-out costs on the order of thousands of dollars to rationalize the fact that, under an opt-in regime, many individuals choose to forego the tax and matching benefits of retirement saving (DellaVigna 2009 and 2018; Bernheim et al., 2015).[3] The required opt-out cost is smaller in my fully-specified dynamic setting because workers can compensate for low initial savings rates by contributing more later. In turn, a smaller opt-out cost implies that the long-term impact of auto-enrollment on asset accumulation may be smaller than previously thought. Indeed, the model predicts that the long-term effect of auto-enrollment on wealth is negligible for a majority of people. However, because the opt-out cost represents a larger share of earnings for low-income individuals, a typical default savings rate of 3% of income adopted by all employers raises retirement wealth at the bottom of the lifetime earnings distribution.[4] In line with the empirical evidence in Chetty et al. (2014) and Beshears et al. (2021), most of the retirement savings increases created by auto-enrollment translate

[3]DellaVigna(2009, 2018), based on a back-of-the-envelope calculation of annual matching benefits, suggest that the minimum opt-out cost to be above $1,000 with time-consistent preferences. Bernheim et al. (2015) estimate an average opt-out cost of around $2,000 (and above $3,000 for the 60% of workers with positive opt-out costs).

[4]In the model the opt-out cost is fixed and identical across individuals. An alternative specification where the opt-out cost is proportional to income does not fit the data well.

into higher total savings with limited crowd-out of other assets and liabilities.

I find support for these long-term predictions in the model's out-of-sample performance in two different settings. This is the main advantage of structural modeling relative to reduced form estimation: the model can extrapolate effects to the long run and to different policies, populations, and institutional environments. First, my model accurately predicts the response of workers in the 86 U.S. 401(k) plans where the default was increased from a positive contribution rate to an even higher rate. Second, the model—with preference parameters estimated in the U.S.—replicates the effect of the U.K. national auto-enrollment policy rollout. The U.S. and U.K. environments vary along multiple dimensions, and the model is calibrated to capture the differences in their retirement savings plans' features, income processes, and tax and benefit systems. The stability of auto-enrollment effects across completely different settings builds confidence in the model's long-term predictions.

Finally, I characterize auto-enrollment's lifetime welfare impact under alternative assumptions about social preferences. If the policymaker shares the same time preferences as individuals and has no redistributive motive beyond declining marginal utility of consumption, an opt-in regime is preferred to auto-enrollment because tax and matching incentives for retirement saving cause individuals to consume more in retirement than implied by their time preferences alone. On the other hand, if the policymaker is paternalistic (i.e. more patient than individuals) or inequality averse (i.e. puts more weight on the welfare of low-income individuals), a default contribution rate near the employer matching threshold maximizes social welfare. My optimal policy characterization is robust to alternative assumptions about the policy incidence; the results hold whether employers reduce their profits, wages, or the match rate to balance their budget in response to the policy.

This paper makes a number of contributions relative to previous literature. While several studies have documented the short-run effects of auto-enrollment on savings (Madrian and Shea, 2001; Choi et al. 2004 and 2006; Beshears et al., 2009; Clark et al., 2015), there are no studies of this policy's long-term effect. This gap in the literature is troubling given that the policy's primary goal is to improve retirement security. In contrast, my paper evaluates the effect of auto-enrollment over a lifetime and studies this policy in a fully-specified dynamic quantitative model.

Another contribution of the paper is to provide a comparison between different behavioral theories that can potentially explain inertia at the savings default. I consider the role of opt-out costs, loss aversion, psychological anchoring, and—in a separate extension—present-biased preferences. My findings complement and broadly align with the results from a randomized control trial by Blumenstock et al. (2018), which found evidence in support of cognitive costs and present bias but did not directly test theories of loss aversion and psychological anchoring.

A number of recent studies have analyzed the welfare impact of auto-enrollment, starting with Carroll et al. (2009), which characterizes the optimal enrollment policy in a model with present-biased preferences, and compares auto-enrollment to a policy forcing an active choice. Bernheim et al. (2015) provides bounds on the welfare effect of different 401(k) default options under several alternative behavioral models, while Bernheim and Gastell (2020) make the case for setting the default at the opt-out minimizing rate under a wide range of distributional assumptions. Goldin and Reck (2020) focus on the normative ambiguity introduced by the fact the opt-out cost may or may not be relevant for welfare. Finally, Zhong (2020) characterizes the optimal default in a setting without employer matching contributions. My paper adds to this literature in two ways: I explore the lifetime welfare impact of auto-enrollment and I consider the policy incidence under alternative means of balancing employers' budget constraints. While the paper makes different assumptions from these previous studies, the resulting optimal policies are quite similar to those identified by Bernheim et al. (2015) and Goldin and Reck (2020): either an opt-in regime or auto-enrollment with a default equal to the employer match threshold. The fact that the optimal policies are similar across these different theoretical frameworks builds confidence in the robustness of the policy recommendations.

The rest of the paper proceeds as follows. In Section 2, I document three stylized facts about auto-enrollment which provide motivation for the model introduced in Section 3. I estimate the model's parameters in Section 4, and explore the sensitivity of my opt-out cost estimate to alternative specifications of the model in Section 5. Section 6 presents results from using the model to predict the long-term effect of auto-enrollment. Section 7 documents the out-of-sample performance of the model. Section 8 analyzes welfare and characterizes optimal policies. Section 9 concludes.

2 Three Stylized Facts about Auto-enrollment

In this section, I document three stylized facts about automatic enrollment which provide motivation for the model. The first fact is that increasing the default contribution rate reduces participation. This fact provides support for a model with a cost of opting out from the saving default and rules out alternative explanations of the default effect. The second fact is that the median cumulative contributions of non-autoenrolled workers catch up to the median cumulative contributions of the autoenrolled workers over three years. Because non-autoenrolled workers are not fully forgoing the benefits of retirement saving by not saving early in their tenure, a relatively small opt-out cost can explain the observed inertia at the default. The third fact is that auto-enrollment reduces savings in subsequent jobs, which suggests that previous estimates have overstated the effect of auto-enrollment on lifetime savings and motivates the need for a lifecycle model to extrapolate the effect of auto-enrollment over many job transitions.

2.1 Data and identification

2.1.1 U.S. 401(k) administrative data

My main identification strategy consists of comparing employees at the same firm hired before and after changes in the default contribution rate.[5] Both groups of workers face the same financial incentives to contribute to the plan, but those hired before the change have a different default option than those hired after.

I compare these two groups of workers using a new proprietary dataset made available by a large U.S. financial institution. It contains 401(k) administrative records from nearly 600 firms with individual-level panel data between December 2006 and December 2017. For each employee, demographic characteristics, participation status in the 401(k) plan, and employee and employer contribution rates (as a share of salary) are recorded annually in December. Contribution dollar amount, 401(k) balances, and asset allocation are recorded at the end of every month. To date, the literature on auto-enrollment in 401(k) plans has relied on either (i) cross-sectional evidence from a large sample of 401(k) plans (Clark et al., 2015), or (ii) within-firm evidence from a small number of firms (Madrian and Shea, 2001; Choi et al. 2004 and 2006; Beshears et al., 2009). My new dataset allows me to combine the strengths of these two approaches by exploiting within-firm variation in the default contribution rate at a large number of 401(k) plans.

2.1.2 U.K. national auto-enrollment policy and representative data

The U.K. Pension Act of 2008 requires employers to automatically enroll their employees (with the option to opt out) into a workplace pension scheme. The law sets the initial minimum employee default contribution rate at 1% of salary and the minimum employer contribution at 1%. The policy was rolled out in phases by employer size between 2012 and 2017, and, unlike the U.S. setting, both new hires and non-participating seasoned employees were automatically enrolled (see Appendix E

[5] A similar identification strategy has been used before by Madrian and Shea (2001) and Choi et al. (2004) using data from four 401(k) plans.

for details on the policy rollout). I exploit the timing of the phased rollout as a source of quasi-experimental variation.

To study this policy, I use individual-level panel data from the U.K. Annual Survey of Hours and Earnings (ASHE) between 2006 to 2017. ASHE is a 1% panel of all U.K. workers, sampled based on the last digits of their National Insurance number. Demographic characteristics, labor earnings, and retirement contributions are collected from administrative tax records and from an annual survey of employers. The data is collected every year in April. There are two advantages to using ASHE. First, it is a nationally representative sample with around 200,000 private sector employees per year. Second, it follows the same individuals over time as they change jobs. However, unlike the U.S. 401(k) administrative data, it does not contain data on workers' retirement account balances and the detailed features of each firms' retirement savings plans.

2.2 Fact I: Increasing the default contribution rate reduces participation

2.2.1 Empirical evidence

I use a sample of 159,216 workers in their first year of tenure in 86 U.S. firms, all of whom were autoenrolled, but with some facing a higher default contribution rate than others.[6] I compare those hired before to those hired after each of these 86 employers increased the default contribution rate from one positive number to a higher one (for instance from 3% to 6% of salary). I estimate the effect of a higher default contribution rate in a linear probability regression with firm fixed effect, year fixed effect, and dummies for each age between 18 and 65. Results are reported in Table 1. The outcome variable in columns 1 and 2 is a dummy equal to 1 if the worker contributes a positive amount to the plan. For each percentage point increase in the auto-enrollment default contribution rate, participation drops by nearly 1 percentage point (columns 1 and 2). In addition, increasing the default contribution rate not only reduces participation, but also increases contributions at the lowest rates. The outcome variable in columns 3 and 4 is a dummy variable for participating in the plan at a contribution rate strictly below the initial default contribution rate. I find that for each percentage point increase in the default contribution rate, there is a 1 percentage point increase in positive contributions below the initial default (columns 3 and 4). For example, when the default increases from 3% to 6% of salary, workers are 3.2 percentage points more likely to contribute either 1% or 2% of their salary to the plan. Finally, columns 5 and 6 measure the combined effect of

[6] I restrict the sample to the first year of tenure to avoid confounding the direct effect of facing a higher default with the indirect wealth effect. Workers enrolled at different default rates accumulate different amount of retirement wealth and, over time, have different incentives to save.

a higher savings default on both participation and lower contributions. Every percentage point increase in the default contribution rate increases the likelihood of contributions below the initial default (including zeroes) by 2 percentage points. In what follows, I show that this new empirical fact can be used to distinguish between alternative explanations of the default effect.

Table 1: The effect of a higher auto-enrollment default on participation and contributions

	Participation (i.e. contribution > 0)		**Positive contribution below initial default**		**Contribution below initial default (inc. 0)**	
	(1)	(2)	(3)	(4)	(5)	(6)
Percentage point increase in the default	-0.008***		0.011***		0.019***	
	(0.001)		(0.002)		(0.003)	
Default increased by 1%		-0.009		0.008***		0.016*
		(0.008)		(0.002)		(0.009)
Default increased by 2%		-0.017***		0.026***		0.043***
		(0.004)		(0.004)		(0.006)
Default increased by 3%		-0.025***		0.032***		0.057***
		(0.007)		(0.010)		(0.016)
Default increased by 4%		-0.036***		0.028***		0.064***
		(0.006)		(0.002)		(0.007)
Employee characteristics	✓	✓	✓	✓	✓	✓
Year FE and Firm FE	✓	✓	✓	✓	✓	✓
Observations	159,216	159,216	159,216	159,216	159,216	159,216
R-squared	0.022	0.022	0.020	0.020	0.020	0.020

* p<0.10, ** p<0.05, *** p<0.01

Notes: Each column reports coefficients from a linear probability regression where the dependent variable is a dummy for the outcome in the column heading. The coefficients correspond to the difference between employees hired before and after each firm increased its auto-enrollment default contribution rate. Each data point corresponds to an employee observed in her first 12 months of tenure, and after the end of the auto-enrollment grace period. Controls for employee characteristics include log of tenure, and dummies for each age between 18 and 65. Standard errors clustered at the firm level are in parentheses. Data source: administrative 401(k) records from 86 U.S. firms.

2.2.2 A theoretical framework for comparing alternative theories

I begin with a simple theoretical framework to compare the predictions of different behavioral theories that can potentially explain inertia at the savings default. This is a simplified version of the full model developed in Section 3.

A worker contributes a fraction s of her wage w to a retirement savings account. She faces a default contribution rate d, and selects an optimal contribution choice s^* from a discrete set S such that:

$$s^* = \underset{s \in S}{\operatorname{argmax}} \left\{ U(s,d) + \delta V_{t+1}(s) \right\}$$

where δ is the intertemporal discount factor, U is flow utility, and V_{t+1} is the continuation value, which is strictly increasing in contributions to the retirement savings account. I use this framework to compare the predictions of three possible explanations of the default effect: opt-out costs, loss aversion and psychological anchoring.

2.2.3 A model of opt-out costs fits the empirical evidence

A first explanation for the default effect is that employees have to incur real or behavioral costs to depart from the savings default. As discussed later in Section 3, these opt-out costs can capture a broad class of neoclassical and behavioral explanations of the default effect such as the cost of completing a form or seeking out advice from a professional adviser, as well as the cognitive cost of paying attention or finding the optimal saving choice.

I assume that workers incur a cost k to choose a contribution rate other than the default:

$$U(s,d) = \begin{cases} u\left((1-s)\,w - k\right) & \text{if } s \neq d \\ u\left((1-s)\,w\right) & \text{if } s = d \end{cases}$$

where the utility function u is increasing and strictly concave in take-home pay $(1-s)\,w$.

Proposition 1 shows that, in line with the empirical evidence, increasing the default contribution rate in a model with a fixed opt-out cost causes more people to drop out completely from the retirement savings plan or to contribute at the lowest rates.[7]

Proposition 1. *If V_{t+1} is concave, increasing the default contribution rate from $\underline{d}$ to $\overline{d}$ weakly increases contributions strictly below $\underline{d}$ in a model with a fixed opt-out cost.*

To build intuition for Proposition 1, consider an employee whose preferred contribution rate is 1% of salary. When the default contribution rate is 3%, this employee may stay at the default to avoid incurring the opt-out cost because the default option is close enough to her preferred contribution rate. However, if the contribution rate is raised to 6%, the same employee may choose to incur the cost and switch to 1% because the higher default is further away from her preferred contribution rate. Therefore, this employee reduces her contribution rate from 3% to 1% when the default is increased from 3% to 6%. A formal proof is provided in Appendix I.[8] Later, in Section 7.1, I show

[7]This pattern of behavior, named the "drop-out" effect" by Caplin and Martin (2017), has been documented in other settings. For instance in the case of charitable giving (Altmann et al., 2018) and tips for taxi rides (Haggag and Paci, 2014), increasing the default led more people to not contribute at all.

[8]The proof relies on the assumption that V_{t+1} is concave in the retirement contribution rate. This assumption

that the opt-out cost model not only makes the right qualitative prediction, but it also predicts the magnitude of employees' response to increasing the default.

2.2.4 Loss aversion makes the opposite prediction

Loss aversion is a common explanation for individuals' propensity to stay at the status quo (Samuelson and Zeckhauser, 1988). Employees may perceive the default as a reference point around which gains and losses are evaluated. For instance, they may perceive an actively chosen additional contribution to the retirement savings plan as a loss in take-home pay, whereas resources that are automatically contributed to the savings account may not feel like losses. To capture this idea, I adopt a functional form similar to DellaVigna et al. (2017) and define flow utility as:

$$U(s,d)=\begin{cases}u\left(\left(1-s\right)w\right)+\eta\left(v\left(s\right)-v\left(d\right)\right) & \text{if } s<d\\ u\left(\left(1-s\right)w\right)+\eta\lambda\left(v\left(s\right)-v\left(d\right)\right) & \text{if } s\geq d\end{cases}$$

where $v(.)$ is decreasing in the contribution rate and reflects gain-loss utility.[9] An employee who contributes below the default contribution rate has more resources than her reference level and experiences gain utility $v(s)-v(d)>0$ with weight $\eta\geq 0$. An employee who contributes above the default has fewer resources than her reference level and experiences loss utility $v(s)-v(d)<0$ with a weight $\eta\lambda$. The parameter $\lambda\geq 1$ captures loss aversion.

Proposition 2 shows that, under loss-averse preferences, increasing the default contribution rate reduces the fraction of employees contributing at the lowest rates.

Proposition 2. *Under loss-averse preferences, increasing the default contribution rate from $\underline{d}$ to $\overline{d}$ weakly decreases contributions strictly below $\underline{d}$:*

$$Pr\left(s^{*}<\underline{d}\,|\,d=\underline{d}\right)\leq Pr\left(s^{*}<\underline{d}\,|\,d=\overline{d}\right)$$

To build intuition for Proposition 2, consider a plan that increased its default contribution rate from 3% to 6% of salary. Contributions at 4%, 5%, or 6% of salary no longer feel like losses. Thus, these contributions become relatively more desirable compared to contributions at 0, 1%, and 2%

may not always hold because of the discrete nature of the contribution choice decision. In practice, if there is enough uncertainty about the future, the continuation value V_{t+1} is generally concave.

[9] A natural candidate functional form is $v=u\left(\left(1-s\right)w\right)$, where take-home pay serves as a consumption reference point. This specification satisfies the condition that v is decreasing in s.

of income (which are in the gain domain under both default contribution rates). A formal proof is provided in appendix I.

2.2.5 Psychological anchoring makes the opposite prediction

People may perceive the default contribution rate as a psychological anchor. Following Bernheim et al. (2015), I assume that the anchoring parameter χ shifts workers preferences toward the value that would rationalize the default as an optimal choice. I assume that workers act as if they were respectively more (less) patient when contributing below (above) the default. The value of contributing s when the default is d is equal to:

$$\begin{cases} u\left(\left(1-s\right)w\right)+\left(\delta+\chi\right)\left(1-m_a\right)V_{t+1}\left(s\right) & \text{if } s<d \\ u\left(\left(1-s\right)w\right)+\delta V_{t+1}\left(s\right) & \text{if } s=d \\ u\left(\left(1-s\right)w\right)+\left(\delta-\chi\right)V_{t+1}\left(s\right) & \text{if } s>d \end{cases}$$

This specification also captures the idea of cognitive dissonance: participants shift their time preference to rationalize why the default may be optimal for them. Similar to loss aversion, Proposition 3 shows that a model of psychological anchoring predicts that increasing the default contribution rate reduces contributions at the lowest rates. A proof is presented in Appendix I.

Proposition 3. *When the default serves as a psychological anchor, increasing the default contribution rate from $\underline{d}$ to $\overline{d}$ weakly decreases contributions strictly below $\underline{d}$:*

$$Pr\left(s^{*}<\underline{d}\,|\,d=\underline{d}\right)\leq Pr\left(s^{*}<\underline{d}\,|\,d=\overline{d}\right)$$

Proposition 3 demonstrates that anchoring contributions at a higher level should always lead employees to contribute and participate more (not less).

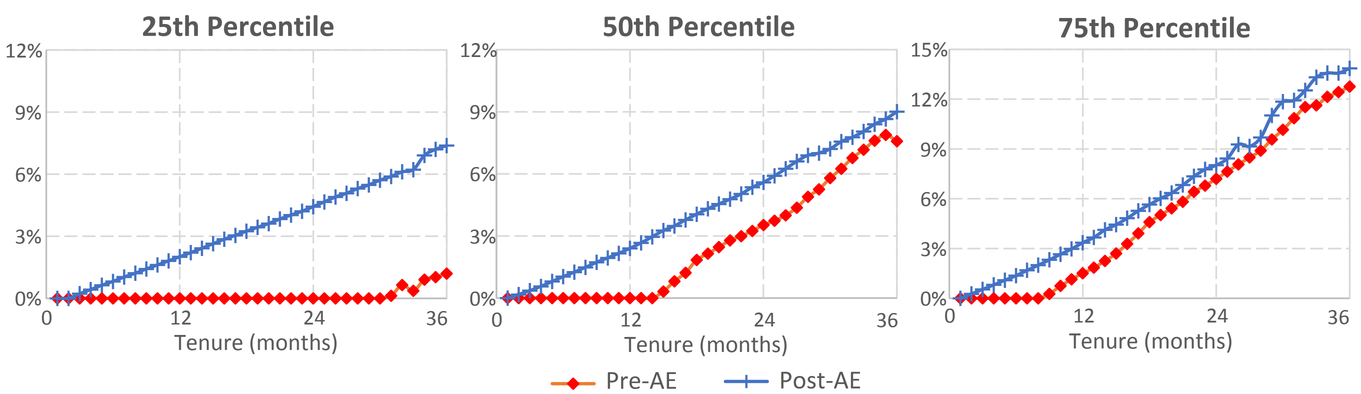
2.3 Fact II: The median cumulative contributions of non-autoenrolled workers catch up to the cumulative contributions of the autoenrolled over 3 years

I document that, at the median, non-autoenrolled workers accumulate the same total amount of retirement savings as autoenrolled workers after 3 years of tenure. While similar evidence has been documented before in smaller samples (Choi et al., 2004),[10] I offer a new interpretation. The fact that

[10] These authors document that the median 401(k) balance-to-pay ratio of non-autoenrolled workers catches up with the autoenrolled after four years in company A, and over 2 years in company B (the same company studied previously

many non-autoenrolled workers catch up with the autoenrolled by contributing more later implies that non-autoenrolled workers are not fully forgoing the match and tax benefits of retirement savings by not contributing early in their tenure. For example, an autoenrolled worker who contributes 3% of salary in her first two years of tenure and a opt-in worker who contributes 0% in her first year and 6% in her second year will earn a similar cumulative employer match and tax benefit over two years. This fact suggests that a relatively modest opt-out cost can generate the observed inertia around the default contribution rate. In turn, a small opt-out cost implies that the long-term effect of auto-enrollment on both wealth accumulation and welfare may be small for a majority of people.

Figure 1: Cumulative employee 401(k) contributions to annual salary ratios at the 25th, 50th and 75th percentiles.



Notes: The graphed series correspond to cumulative employee 401(k) contributions at each month of tenure divided by individuals' annual salary. The Pre-AE (Post-AE) cohort corresponds to employees hired in the 12 months before (after) the introduction of an auto-enrollment policy at 3% of salary for new hires. The sample at each tenure level corresponds to workers still employed by the firm at that time. Data source: administrative 401(k) records from 34 U.S. firms.

The empirical evidence. I compare employees hired in the 12 months before to employees hired in the 12 months after the introduction of auto-enrollment in 34 U.S. firms. These firms offer a 50% match on contributions up to 6%, and moved from an opt-in to an auto-enrollment regime at 3% of salary for new hires (this sample is described in more detail in Section 4.2). This match formula and auto-enrollment default are the most common both in my sample of 401(k) plan and in nationally-representative data (Arnoud et al., 2021). Figure 1 shows the distribution of cumulative employee contributions in the two groups. The numerator is the sum of all contributions made by an employee to the retirement savings plan since her date of hire, and the denominator is her annual salary. Autoenrolled workers initially contribute more to the retirement savings plan. However, both at the median and at the 75th percentile, the non-autoenrolled employees catch up over three years

by Madrian and Shea, 2001).

and contribute a similar total amount to the 401(k) plan as the autoenrolled workers (Figure 1 middle and right panels). At the bottom of the distribution of cumulative employee contributions, however, the savings increases created by auto-enrollment persist over three years (Figure 1 left panel). The catching-up behavior observed in the opt-in group is very similar when looking at employees hired more than 12 months prior to the policy suggesting that it is not driven by peer effects.[11]

Why previous studies estimated implausibly large opt-out costs. The middle panel of Figure 1 documents that, at the end of the first year of tenure, median cumulative employee contributions are equal to zero for non-autoenrolled workers and equal to around 3% of salary for autoenrolled workers (i.e. the auto-enrollment default in the sample). This implies that, by the end of their first year of tenure, autoenrolled workers have earned substantially more employer matching and tax benefits than non-autoenrolled workers. Because the difference in benefits earned by the autoenrolled and non-autoenrolled workers is large, previous studies have suggested that implausibly large opt-out costs are required to rationalize inertia around the default contribution rate. DellaVigna (2009, 2018) sets the minimum opt-out cost above \$1,000 with time-consistent preferences, based on a back-of-the-envelope calculation of annual matching benefits for the median worker. Bernheim et al. (2015) estimate a median opt-out cost of \$2,200 using data on workers' first few quarters of tenure. These opt-out cost estimates come from settings that do not allow for future adjustment of the contribution rate and are therefore sensitive to the period of observation.

Why the opt-out cost may be smaller in a fully dynamic setting. By the end of the third year, median cumulative contributions of non-autoenrolled workers catch up to those of autoenrolled workers (Figure 1 middle panel). This fact implies that both groups of workers earn a similar total amount of employer matching contribution and tax benefits over three years.[12] Hence, applying the previous back-of-the-envelope calculation using data from the end of the third year of tenure would imply that the opt-out cost is negligible at the median. Because the estimated size of the opt-out costs is sensitive to the period of observation in models that do not allow for the dynamic adjustment of contributions, it is necessary to use a fully dynamic model to obtain an estimate of the opt-out cost that is consistent across observation periods. I show in Section 4.7 that in a dynamic lifecycle

[11]Because employees hired in the 12 months prior to the policy interact with an increasing number of auto-enrolled new colleagues, their catching-up behavior could be partly driven by peer effects. However, the evolution of participation and contributions over three years is very similar when looking at employees hired respectively between 12 and 24 months, 24 and 36 months, or 36 and 60 months prior to the policy. These prior cohorts did not overlap as much (or at all) with the auto-enrollment cohort and, therefore, could not have been subject to their peer effect in their first three years of tenure. These figures are available on request.

[12]Note that, at the median, neither the autoenrolled nor the non-autoenrolled workers contribute above the 6% threshold on employer matching. Thus, a similar level of cumulative employee contributions implies a similar level of cumulative employer contributions.

model an opt-out cost of around $250 generates the patterns of behavior documented in Figure 1.

Implication for auto-enrollment's long-term effect on wealth. Given that, at the median and above, non-autoenrolled workers' cumulative contributions catch up over three years with the autoenrolled's cumulative contributions (Figure 1), the long-term auto-enrollment effect on wealth is likely to be small for a majority of people. However, savings increases created by auto-enrollment persist at the 25th percentile of the distribution of cumulative employee contributions. Will the large effect on retirement savings at the bottom persist in the long run and create welfare gains that justify the auto-enrollment policy? Since many of these employees would have left the employer after a few years, the answer to this question depends on how autoenrolled employees behave after they leave their current employers. In the next subsection, I study this question in the context of the U.K. national auto-enrollment policy.

2.4 Fact III: Current auto-enrollment reduces saving in the subsequent job

I exploit the phased roll-out of the U.K. national auto-enrollment policy to study the participation and contributions of employees moving from a firm subject to the auto-enrollment requirement (larger firms) to a firm not yet affected by the policy (smaller firms). I compare their saving decisions in their first year of tenure to the saving decisions of workers who moved to the same firm in the same year, but coming from an employer not yet affected by the policy (i.e. from a smaller to another smaller firm).

I use data from 12 successive waves of ASHE between April 2006 to April 2017. The policy rollout started in 2012 with only the largest employers affected, and in every ASHE wave between 2012 and 2017 I observe more firms subject to the auto-enrollment requirement (see Appendix E for details of the rollout). By the last wave in April 2017, all employers with more than 30 employees were subject to the auto-enrollment mandate. I include data prior to the beginning of the policy rollout to control for time-invariant differences between workers moving between firms of different sizes. The sample contains 37,120 job-switchers in their first year of tenure at their new employer.

I estimate the following specification for a worker i who transitioned in year t from employer e_{n-1} to employer e_n:

$$y_{i,t} = \beta_0\, AEtoNoAE_{i,t} + \beta_1\, AEtoAE_{i,t} + \sum_{j=1}^{7} \phi_j\, Size_j^{e_{n-1}} + \sum_{j=1}^{7}\sum_{g=1}^{7} \gamma_{j,g}\, Size_j^{e_{n-1}} \times Size_g^{e_n} + \varphi\, Z_{i,t} + \alpha_{e_n,t} + \varepsilon_{i,t}$$

I consider two outcome variables $y_{i,t}$, employees' participation status in the retirement savings plan and their contribution rate as a share of salary. The variable $AEtoNoAE_{i,t}$ is a dummy equal to 1 for worker i in year t if both: (i) her current employer e_n is not subject to the auto-enrollment requirement in year t, and (ii) her previous employer e_{n-1} was subject to the auto-enrollment requirement the last time employee i was observed at firm e_{n-1}. In the presence of firm-year fixed effects $\alpha_{e_n,t}$, the coefficient β_0 captures the difference in participation and contribution rates among new hires in the same opt-in firm who were or were not subject to the auto-enrollment requirement in their previous job. Similarly, the variable $AEtoAE_{i,t}$ is a dummy equal to 1 if the worker were subject to automatic enrollment both in her current and in her previous job, and the coefficient β_1 captures the difference in participation and contribution rates among new hires in the same auto-enrollment firm who were or were not subject to the auto-enrollment requirement in their previous job. Since the policy was rolled out by the size of employers, one concern could be that new hires coming to firm e_n from larger previous employers (which were subject to auto-enrollment earlier) are systematically different from those joining from smaller ones. I use data before and after the policy rollout to identify the auto-enrollment effect separately from the effect of moving from a larger to smaller employer. To do so, I control for a set of interaction dummies $Size_{e_{n-1}} \times Size_{e_n}$ for each pair of current and previous employer size categories (i.e. with 7 firm size categories I have 49 pairs of current and previous employer sizes).[13] I use data prior to the beginning of the auto-enrollment rollout to identify these interaction dummies separately from the effect of moving from a non-autoenrollment to an auto-enrollment employer. Finally, $Z_{i,t}$ is a set of individual controls (a third order polynomial in age, gender, log salary and log tenure in the current job e_n, and log salary and log tenure when last observed in the previous job e_{n-1}).

My baseline estimate for β_0 is reported in the first column of Table 2. Workers who were previously auto-enrolled participate 12.8 percentage points less and contribute 0.55% of salary less to the retirement savings plan of their next non-autoenrollment employer. In contrast, when workers are not required to make an active decision in their new job (i.e. moving to a new employer subject to the automatic enrollment mandate) there is no statistically significant difference between those who were previously auto-enrolled and those who weren't. To test that these estimates capture the effect of the policy rollout and not differences across employers, I conduct a number of placebo exercises. I run the same specification under the assumption that the policy rollout started before the actual policy implementation date in 2012. These placebo tests, reported in columns 2 to 8 of

[13] These seven employer size categories are firms with more than $30,000$ workers, between $6,000$ and $29,999$ workers, 350 to $5,999$ workers, 160 to 349 workers, 58 to 159 workers, 50 to 57 workers, and 30 to 49 workers. Firms with fewer than 30 employees is the omitted category.

Table 2, are all insignificant at the 5% level for both participation and contribution outcomes, and a single one is significant at the 10% level (column 7). These results help validate the empirical strategy and suggest that the estimates capture the effect of the auto-enrollment rollout rather than the effect of moving from a larger to smaller employer.

Table 2: Auto-enrollment effect after a job transition

	Actual policy	**Placebo tests**						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Beginning of policy rollout	2012	2005	2006	2007	2008	2009	2010	2011
			Panel A - Participation rate					
AE to non-AE employer	-0.128**	0.077	0.034	0.014	0.016	0.053	0.016	-0.061
	(0.055)	(0.058)	(0.037)	(0.051)	(0.049)	(0.058)	(0.051)	(0.086)
AE to AE employer	0.010	0.001	0.007	0.014	0.008	0.016	0.028*	0.021
	(0.019)	(0.017)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)
			Panel B - Contribution rate (in percentage of pay)					
AE to non-AE employer	-0.546**	-0.001	-0.018	0.261	-0.139	0.003	-0.269	-0.209
	(0.226)	(0.209)	(0.187)	(0.541)	(0.221)	(0.244)	(0.217)	(0.336)
AE to AE employer	-0.039	-0.013	-0.026	-0.04	-0.001	0.052	-0.026	-0.035
	(0.068)	(0.093)	(0.073)	(0.064)	(0.062)	(0.068)	(0.099)	(0.069)
Employee characteristics	✓	✓	✓	✓	✓	✓	✓	✓
$\text{Size}_{e'} \times \text{Size}_e$	✓	✓	✓	✓	✓	✓	✓	✓
Employer×Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	37120	37120	37120	37120	37120	37120	37120	37120

* p<0.10, ** p<0.05, *** p<0.01

Notes: Each cell reports coefficients from a regression whose dependent variable is a dummy for workers' participation status in the retirement savings plan (Panel A), or contribution rate in percentage points of pensionable pay (Panel B). The sample contains job switchers in their first year of tenure. The coefficients correspond to the effect of the automatic enrollment mandate in the previous employer on contributions when the current employer is not subject to the auto-enrollment requirement (AE to non-AE employer) and when the current employer is subject to automatic enrollment (AE to AE employer). In Column 1, I assign an enrollment regime to each employer (i.e. opt AE or non-AE) based on the U.K. national auto-enrollment policy rollout schedule (in Appendix E). In Columns 2 to 8 I assign enrollment regimes based on a (placebo) policy rollout schedule that starts earlier than the actual policy. I assume that the policy rollout started in the year reported in the column heading. Controls for employees' characteristics include controls for log salary and log tenure in the current job, log salary and log tenure when last observed in the previous job, a dummy for gender, and a third order polynomial in age. The sample is restricted to private sector workers between the ages of 22 and 60. Standard errors clustered by current employer are in parentheses. Data source: U.K. Annual Survey of Hours and Earnings waves 2006 to 2017.

My findings go against the hypothesis that auto-enrollment created lasting saving habits, in which case it should have caused workers to contribute more in their next job's retirement savings plan. If anything, when required to make an active enrollment decision in their next job, previously auto-enrolled workers are less like to contribute to the new plan. There are several possible explanations for this finding: individuals who were induced to accumulate more savings in their previous job could have less incentive to contribute later on because of a standard wealth effect. Alternatively, previously auto-enrolled individuals may choose not to opt in their new employer's retirement plan because they are better informed about the future rollout of the policy and expect that their new employer will soon enroll them automatically.[14] I discuss and assess these alternative explanations in Appendix E. While I cannot isolate the mechanisms behind the results, I show in Section 7.2 that the magnitude of the estimates for both types of transitions (AE to nonAE and AE to AE) is consistent with the model-implied size of the wealth effect. Auto-enrolled individuals accumulate more retirement wealth early-on and, therefore, have less incentive to contribute to their next job's retirement savings plan.[15]

On average, U.S. individuals change jobs more than ten times over their lifetime (BLS 2019). In the next section, I build and estimate a structural lifecycle model to extrapolate the effect of auto-enrollment across many job transitions. Motivated by facts I and II, the model features a cost of opting out of the default and allows employees to adjust their contribution rate over time.

3 A Lifecycle Model with Default Effects

Motivated by the empirical evidence presented in the previous section, I develop a new lifecycle model with default effects. Before detailing the model, I provide a short summary of its key features. Every period, individuals choose a level of consumption, retirement wealth, liquid wealth, and unsecured debt. Individuals incur a utility cost every time they choose a retirement contribution rate that is different from the default. Their contribution choice becomes, in turn, the next period's default's contribution rate. Other important components of the model are: i) a labor market with uninsurable earnings and employment risk; ii) a progressive tax and benefits system with Social Security-style public pensions, iii) realistic retirement savings accounts with employer matching contributions and favorable tax treatment, and iv) demographic change over the lifecycle with stochastic mortality and evolving household composition.

[14] Note that if all new hires have rational expectations about the future policy rollout, we should not expect those who were previously autoenrolled and those who weren't to have different expectations about the future roll-out of the policy.

[15] Interestingly, the model prediction is closer to the empirical estimate for job switchers whose employer will not become subject to the auto-enrollment mandate in the following 12 months (and who are therefore less subject to the confounding effect of expectations about an imminent adoption of auto-enrollment).

In what follows, I introduce the baseline model. Table 6 in Appendix A gives a summary of all the parameters introduced in this section. In Appendix B, I explore three extensions to the baseline model with respectively present-biased preferences, an opt-out cost proportional to employees' earnings, and no offset of unemployment benefits by retirement withdrawals.

3.1 Opt-out costs

Individuals in the model act as if they are required to pay a utility cost every time they change their contribution rate. This opt-out cost captures both real transaction costs as well as several behavioral explanations of the default effect, such as the cost of seeking advice from a financial adviser or the cognitive costs of paying attention or finding an optimal choice. A broad class of behavioral theories, axiomatized by Masatlioglu and Ok (2005), satisfies this representation. Goldin and Reck (2020) show a correspondence between these opt-out costs and models of inattention, status-quo bias and defaults as implicit advice. Consistent with the importance of cognitive and complexity costs, both Choi et al. (2009) and Blumenstock et al. (2018) find that more employees make an active decision when the enrollment process is simplified or when they receive help computing the implications of alternative savings scenarios.

The opt-out cost model also captures the role of present bias. While present bias does not, by itself, generate inertia around the default option, it can exacerbate the effect of a small opt-out cost. People with present-biased preferences behave as if the opt-out cost is larger because they incur this cost in the present while the benefits of retirement saving are in the future (O'Donoghue and Rabin, 1998; Carroll et al., 2009; DellaVigna 2009 and 2018). My baseline opt-out cost estimate can thus be interpreted as capturing the combined effect of present bias and other real and behavioral costs.[16] In a separate extension in Appendix B.1, I reject a model where present bias directly affects contributions.[17]

I assume that the opt-out cost is the same for all individuals and in all periods. While Bernheim et al. (2015) allow for heterogeneous opt-out cost and each of Carroll et al. (2009) and DellaVigna (2018) consider stochastic and time-varying costs, I show that my more parsimonious specification can endogenously generate the heterogeneity observed in the data along each of age, income, and

[16] Note that it is, in principle, possible to identify present bias separately from other forms of contribution switching costs using data on both 401(k) contributions and liquid asset holdings. Beshears et al. (2021) study a setting with such linked data.

[17] Contribution decisions are implemented by employers with a delay (usually at the beginning of the next pay period). As noted by Carroll et al. (2009), this implies that a contribution rate choice may serve as a commitment to save in the future, starting from the next paycheck, which means that the contribution rate choice is unaffected by present bias.

tenure dimensions (in Section 4) and accurately predict behavior out-of-sample (in Section 7).[18]

3.2 Demographics and preferences

Age is indexed by a and working life starts at age 22. Each period t is one quarter. Mortality is stochastic; individuals survive each period with probability $1-m_a$ and die with certainty at age A. The period utility function, given in equation (1), is of the isoelastic form with a constant elasticity of intertemporal substitution σ. An equivalence scale n_a in the utility function captures the evolution of household composition over the lifecycle.

$$u_a(\cdot) = n_a \frac{\left(\frac{\cdot}{n_a}\right)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \tag{1}$$

3.3 Labor market

Individuals in the model face uninsurable earnings and employment risk. Every period, they can be in one of four employment states emp_t: continuously employed (E), newly employed following a job-to-job transition (JJ), unemployed (U), or retired (Ret).

Continuous employment. When employed, individuals receive uncertain labor earnings. The log of labor earnings, given in equation (2), is equal to the sum of a deterministic component (a cubic in age) and a stochastic component θ_t.

$$\ln w_t = \delta_0 + \delta_1 a_t + \delta_2 a_t^2 + \delta_3 a_t^3 + \theta_t \tag{2}$$

The stochastic component θ_t follows an AR(1) process with normally distributed innovations ξ (equation (3)). The variance of these innovations is different in the first period than in subsequent periods to account for initial heterogeneity in income at age 22.

$$\theta_t = \rho\theta_{t-1} + \xi_t \quad \text{if}\, emp_t = E \tag{3}$$

$$\xi_1 \sim \mathcal{N}\left(0, \sigma_{\xi_1}^2\right)$$
$$\xi_t \sim \mathcal{N}\left(0, \sigma_{\xi}^2\right)\ \forall\, a > 1$$

[18]Given that my specification with a homogeneous opt-out cost fits well the dimensions of heterogeneity that are observable in the data, there isn't sufficient empirical variation left to credibly identify the heterogeneity in opt-out costs. In principle, future work could exploit linked administrative and survey data (similar to the setting of Goda et al., 2020) to identify the full distribution of opt-out costs.

Job-to-job transitions. Job-to-job transitions are modeled separately from continuous employment because pension arrangements change when moving jobs. An employee switches employer with a probability $\pi^{JJ}(a, ten)$ that depends on her age and her tenure in her current job. A job-to-job transition is associated with a new earnings draw (equation (4)). The labor earnings innovation ζ_t has mean $\mu^{JJ} > 0$ to capture the earnings premium associated with job-to-job transitions.

$$\theta_t = \rho\theta_{t-1} + \zeta_t \quad \text{if } emp_t = JJ \tag{4}$$

$$\zeta_t \sim \mathcal{N}\left(\mu^{JJ}, \sigma_\xi^2\right)$$

Unemployment. The probability of moving into unemployment $\pi^{EU}(a, ten)$ depends on both age and tenure. Unemployed individuals receive unemployment insurance benefits $ui(\theta_t)$, described in Section 3.7. The stochastic component of earnings θ_t stays constant during an unemployment spell. An unemployed individual transitions back to employment with an age-dependent probability $\pi^{UE}(a)$. The earnings innovation in the new job ς_t has mean $\mu^{UE} < 0$ to capture the earnings loss associated with job displacement (equation (5)).

$$\theta_t = \rho\theta_{t-1} + \varsigma_t \quad \text{if } emp_t = U \tag{5}$$

$$\varsigma_t \sim \mathcal{N}\left(\mu^{UE}, \sigma_\xi^2\right)$$

Retirement. Individuals retire deterministically at age A^{ret}. They receive a Social Security-style public pension in retirement, described in Section 3.7.

3.4 Decisions

In the model, people save for both precautionary motives (to insure against bad labor earnings realizations and unemployment risk) and for life-cycle motives (to complement Social Security in retirement). Every period, they make two decisions: how much to hold in liquid taxable wealth (l_t) and how much to contribute or withdraw from a tax-deferred defined contribution (DC) retirement account (dc_t).

Liquid taxable wealth. People start their lives with zero liquid wealth. Every period thereafter, they choose a level of liquid asset holdings l_{t+1}. Contributions to the liquid asset are made out of after-tax income. Positive liquid asset holdings grow at a constant rate $1 + r$ and these returns are subject to a capital tax τ^k. Borrowing up to a limit $\underline{l}_a$ is allowed, reflecting the availability of unsecured credit. Borrowing is subject to a higher interest rate $1 + r^{CC}$. The after-tax interest rate on liquid assets is thus given by:

$$R^{liq}(l) = \begin{cases} 1 + r\left(1 - \tau^k\right) & if\, l \geq 0 \\ 1 + r^{cc} & if\, l < 0 \end{cases}$$

DC contributions. Employees can contribute a fixed percentage of their labor earnings $s_t \geq 0$ to a DC account. Contributions are subject to an age-dependent maximum contribution limit set by the government, such that $w_t s_t \leq limit_a$. DC contributions are tax-deferred and, thus, reduce taxable income in the contribution period. DC wealth dc_t earns the same risk-less return r as the liquid asset but these returns are not subject to capital taxation ($R^{dc} = 1 + r$).

DC withdrawals. People can withdraw a fraction $draw_t \in [0, 1]$ of their DC wealth when unemployed or retired.[19] Withdrawals are subject to both income taxation and—depending on the individual's age—an additional tax penalty pen_a for early withdrawals.

3.5 Retirement savings plans

Employers are heterogeneous in the matching formula and default option of their retirement savings plan (indexed by e). Employer contributions to the DC account, denoted ec^e, are characterized by a level of non-contingent contributions ($fixed^e$), a match rate that is a function of tenure ($match^e(ten)$), and a threshold contribution rate for matching contributions (cap^e). Contributions above this threshold do not earn an employer match contribution. In some plans, an employee who separates from her employer before the end of a vesting period may lose part (or all) of the employer matching contributions. To capture this effect without introducing an additional state variable, I assume that the employer match is adjusted by a factor $\Upsilon_e(a, ten) \leq 1$ that depends both on the vesting schedule and on the probability that an employee—based on her age and tenure—separates from her employer before the end of the vesting period (see Appendix D.1.2 for more details). An

[19] Some 401(k) plans offer hardship withdrawals and loan options for employed participants, but these early distribution options are discouraged by the tax code. During the sample period, hardship withdrawals are subject to a 10% tax penalty and trigger a minimum six months suspension from contributing to the plan. 401(k) loans are only exempt from the tax penalty if repaid in full before the worker leaves her job and interest payments are taxed twice: payments are made with after-tax dollars and are taxed again at distribution (see Beshears et al., 2008).

employee contributing s_t receives an employer DC contribution equal to:

$$ec^e\left(s_t, a_t, ten_t\right) = \left(fixed^e + \Upsilon_e\left(a_t, ten_t\right) \times match^e\left(ten_t\right) \times \min\left\{s_t, cap^e\right\}\right) w_t$$

3.6 Default mechanism

Default contribution rate. In the first period in a job, the default contribution rate $\bar{d}^e$ is drawn from the exogenous distribution of employer types $E\left(\boldsymbol{.}\right)$. It is equal to zero under the opt-in enrollment regime and strictly positive under auto-enrollment. Later in a worker's tenure, the default equals her contribution rate in the previous period; it is therefore endogenous to individuals' saving decisions:

$$d_t = \begin{cases} \bar{d}^e \sim E\left(\boldsymbol{.}\right) & \text{if } ten_t = 1 \\ s_{t-1} & \text{if } ten_t > 1 \end{cases}$$

Opt-out cost. In order to change their contribution rate away from the default, people have to incur a utility cost k. The period-utility of an employee contributing s_t to her DC plan is given by:

$$u_a\left(c_t - \mathbb{1}_{(s_t \neq d_t)} k\right)$$

As discussed previously in Section 3.1, this utility cost k captures both neoclassical transaction costs as well as a wide range of behavioral costs such as the cognitive cost of paying attention or finding the optimal choice.

3.7 Government

Taxes. Agents face a non-linear income tax schedule $tax^i\left(\boldsymbol{.}\right)$. DC accounts are treated favorably by the tax system: contributions are not subject to income taxation, but withdrawals (in either unemployment or retirement) increase taxable income by the withdrawal amount.[20] Taxable income is equal to labor earnings plus DC withdrawals minus DC contribution. Returns on capital are taxed at a constant rate τ^k but DC wealth is exempt from capital taxation.

Unemployment benefits. People receive a benefit $ui\left(\theta_t\right)$ when unemployed. This benefit

[20] This is the traditional tax-deferred DC model. I abstract from the more recent, and less common, Roth-401(k) tax model where DC contributions are taxed up front.

depends on labor productivity in the last period of previous employment. In the baseline model, the employer contribution protion of a withdrawals from the retirement account is treated as compensation and may offset unemployment benefits (details in Appendix D.1.2).

Public pension benefits. After retirement, individuals receive Social Security-style public pension benefits. Public pension benefits $ss(\cdot)$ are computed according to a non-linear schedule with an income floor. The amount received depends on the individual's average lifetime earnings ae_{T^R}. The law of motion of average lifetime earnings is given in equation (6).

$$ae_{t+1} = \begin{cases} \frac{w_{t+1}+a_t ae_t}{1+a_t} & \text{if } a_t < A^{ret} \\ ae_{A^{ret}} & \text{if } a_t \geq A^{ret} \end{cases} \tag{6}$$

3.8 State variables

The dynamic optimization problem is characterized by 9 state variables: age (a), employment status (emp), tenure (ten_t), employer DC plan type (e), labor productivity (θ), average lifetime earnings (ae), liquid assets (l), DC wealth (dc) and the default contribution rate d. The vector of state variables is denoted $X_t = [a_t, emp_t, ten_t, e_t, \theta_t, ae_t, l_t, dc_t, d]$. There is uncertainty over survival to the next period (with mortality probability (m_a)), the stochastic component of earnings (θ), employment status (emp), and the type of the future employer after a job change (e). The joint distribution of earnings, employment and employer types is denoted $F(\theta_t, emp_t, e_t)$.

3.9 Individual maximization problem

Individuals maximize discounted intertemporal utility with an exponential discount factor δ. In what follows, I introduce the recursive optimization problem in periods of retirement, employment, and unemployment.

Individual problem in retirement (after the age of 65). The only source of uncertainty in retirement is mortality risk. Retired individuals decide every period how much to hold in the liquid asset and how much to withdraw from the DC account.

$$V_t(X_t) = \max_{draw_t \in [0,1],\, l_{t+1} \geq \underline{l}_a} u_a(c_t) + \delta(1-m_a) V_{t+1}(X_{t+1})$$

$$s.t.\ l_{t+1} = R^{liq}(l_{t+1}) \left[ss(ae_{A^{ret}}) + draw_t dc_t + l_t - c_t - tax_t^i \right]$$

$$dc_{t+1} = R^{dc}((1-draw_t) dc_t)$$

Individual problem when employed. People are employed following a job-to-job transition, an unemployment-to-employment transition or while in continuous employment. They face earnings and employment risk and decide how much to save in liquid wealth or borrow through unsecured credit and how much to contribute to the DC account. The decision is in two steps.

In the first step, for every possible DC contribution choice s in the discrete set S, the worker solves for the optimal amount of liquid wealth or unsecured debt to hold subject to the intertemporal budget constraint 7. The value of choosing a contribution rate s is thus given by:

$$V_t^s(X_t) = \max_{l_{t+1}} u_a\left(c_t - \mathbb{1}_{(s_t \neq d_t)} k\right) + \delta(1-m_a) \int V_{t+1}(X_{t+1})\, dF(\theta_t, emp_t, e_t)$$

$$s.t.\ l_{t+1} = R^{liq}(l_{t+1})\left[(1-s_t) w_t + l_t - c_t - tax_t^i\right] \tag{7}$$

$$l_{t+1} \geq \underline{l}_a$$

In the second step, the worker selects a DC contribution rate s subject to a tax limit on DC contributions:

$$V_t(X_t) = \max_{\mathrm{s_t \in S}} \left\{V_t^s(X_{t+1})\right\}$$

$$s.t.\ s_t w_t \leq limit_a$$

$$dc_{t+1} = R^{dc}\left(dc_t + s_t w_t + ec_t^e\right)$$

Individual problem when unemployed. After an unemployment shock, people receive constant unemployment benefits and face uncertainty about finding a job and the wage and type of DC plan in their future job. They decide how much liquid wealth or unsecured debt to hold and how much to withdraw from the DC account.

$$V_t(X_t) = \max_{draw_t \in [0,1],\, l_{t+1} \geq \underline{l}_a} u_a(c_t) + \delta(1-m_a) \int V_{t+1}(X_{t+1})\, dF(\theta_t, emp_t, e_t)$$

$$s.t.\ l_{t+1} = R^{liq}(l_{t+1})\left[ui_t + draw_t(1-pen_a)\, dc_t + l_t - c_t - tax_t^i\right]$$

$$dc_{t+1} = R^{dc}(1-draw_t)\, dc_t$$

This dynamic optimization problem does not admit a closed form analytical solution. The numerical solution procedure is discussed in appendix G

4 Estimation and Model Fit

4.1 Method of Simulated Moments

I estimate the model parameters in two stages.[21] In the first stage, I set a number of parameters outside of the model, either estimated directly from the data or coming from previous literature. In the second stage, I use the Method of Simulated Moments to estimate the model's three preference parameters: the intertemporal discount factor (δ), the elasticity of intertemporal substitution (σ), and the opt-out cost (k).

4.2 Data and quasi-experimental variation

Estimation sample. I estimate the model's preference parameters using individual-level 401(k) contribution data on employees hired right before and right after the introduction of auto-enrollment. I use the dataset described previously in Section 2.1.1, and I restrict the sample to 401(k) plans that: (i) have a 3% initial auto-enrollment default contribution rate and no auto-escalation feature,[22] (ii) offer a 50% employer match contribution up to 6% of income, (iii) and for which the exact date of auto-enrollment adoption is available. 34 firms satisfy these inclusion criteria. I restrict the sample to employees hired in these 34 firms in the 12 months before or 12 months after the introduction of automatic enrollment, and who are observed after the end of the auto-enrollment grace period (3,682 employees before and 2,733 after the policy).[23] I use this sample to estimate second-stage parameters.

Quasi-experimental variation. Figure 2 shows the frequency of contributions at exactly 3% of salary—the default contribution rate—among employees with up to 12 months of tenure, by date of hire relative to the date of the adoption of auto-enrollment. In their first year of tenure, employees hired under auto-enrollment are 70 percentage point more likely to contribute the default contribution rate (i.e. exactly 3% of salary) than employees hired in the year before the introduction of the policy. The magnitude of these results is in line with previous evidence in Choi et al. (2006) from 4 firms using the same identification strategy. I exploit this quasi-experimental variation in the enrollment regime to identify the model preference parameters.

[21] This two-step procedure is typical in the estimation of lifecycle models (see, among others, Gourinchas and Parker (2002), Low et al. (2010), Laibson et al. (2017), and O'Dea (2018)).

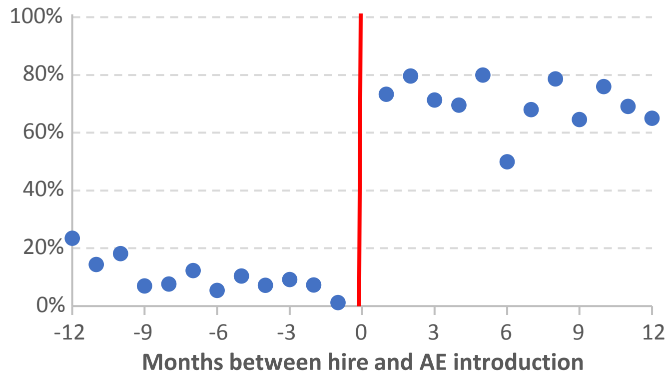
[22] Auto-escalation is a feature in some retirement plans, whereby the contribution rate increases automatically every year unless the participants actively opt out.

[23] Employees can withdraw, without a tax penalty, any amount automatically contributed to a 401(k) plan if they opt out before the end of a grace period, typically one to three months from the date of hire. Workers forfeit any matching contribution that would have been made for the automatic enrollment contributions.

4.3 Identification of preference parameters

The three model preference parameters are jointly estimated. In what follows, I discuss which variation in the data allows for the identification of each parameter.

Figure 2: Share of workers contributing the default contribution rate in their first year of tenure



Notes: Every point corresponds to the share of workers, with up to 12 months of tenure, contributing exactly 3% of salary. The horizontal axis corresponds to the difference between workers' date of hire and the date automatic enrollment was introduced for new hires. Data source: administrative 401(k) records from 34 U.S. firms.

Time preferences. The model has two parameters governing saving preferences: the elasticity of intertemporal substitution (σ) and the intertemporal discount factor (δ). Since both parameters affect saving decisions, it is often difficult to obtain separate identification from lifecycle saving profiles alone. To overcome this challenge, I exploit a discontinuity in the savings interest rate at the employer matching contribution threshold. In the estimation sample, contributions up to 6% of income are matched by the employer at 50%, while contributions above the threshold are not. Bunching of employee contributions at 6% can therefore be used to identify the elasticity of intertemporal substitution separately from the discount factor, which governs the overall level of saving. This approach is similar to Best et al. (2018), who use jumps in the mortgage interest rate at thresholds for the loan-to-value ratio to identify the elasticity of intertemporal substitution.

Opt-out cost. I compare the contribution behavior of employees hired before versus after the adoption of auto-enrollment to identify the size of the opt-out cost. Since both groups of employees face similar economic incentives, absent the opt-out cost, they should make similar contribution choices. Bunching of contributions at the 3% default contribution rate in the auto-enrollment group (and its evolution with tenure) identifies the size of the opt-out cost. This identificaiton strategy is similar to that of Bernheim et al. (2015): we should expect no bunching at the default if the opt-out cost is zero, and 100% bunching at the default if the opt-out cost is infinitely large.

4.4 First-stage parameters

I use data from the Survey of Income Programs and Participation (SIPP) to estimate the parameters of the labor earnings process and labor market transition probabilities at the quarterly level. The remaining first-stage parameters (including demographic variables, the parameters of the tax and benefits system, the vesting schedule, and the rate of return on assets) are estimated directly from the data or calibrated with reference to other papers. The estimation of the earnings process, treatment of the data, and calibration of other first-stage parameters are discussed in Appendix D.

4.5 Second-stage parameter estimation

The three second-stage parameters (δ, σ, and k) are estimated using the Method of Simulated Moments. The parameter estimates are chosen to minimize the distance between model-simulated and empirical contribution patterns for employees hired before and after the introduction of auto-enrollment.

Simulation. To perform the estimation, I simulate $5,000$ individuals under two scenarios. In the opt-in scenario, individuals face a 0% default contribution at the start of every job. In the auto-enrollment counterfactual, they experience the same sequence of employment and productivity draws. However, each simulated individual—at one randomly drawn job transition—*unexpectedly* faces a 3% default contribution rate when starting a new job. I compare simulated contribution behavior under these two scenarios to the empirical patterns of employees hired before and after the introduction of auto-enrollment.

Estimation moments. I use two sets of moment conditions (44 in total). First, I use the distribution of contribution rates among employees in their first year of tenure: the fraction of workers contributing 0, 3%, 6%, and 10% or above of salary under both the opt-in and the auto-enrollment regimes (8 moment conditions). Second, I use the evolution of contributions and participation over the first 4 years of tenure: participation rate under opt-in, participation rate under auto-enrollment, and contributions at the auto-enrollment default contribution rate (i.e. exactly 3% of salary) every quarter between the 5^{th} and 16^{th} quarters of tenure (36 moment conditions).

Weighting matrix. In my baseline specification, I use as a weighting matrix the inverse of the diagonal of the variance-covariance matrix of the second-stage moment conditions obtained by bootstrap. I use this diagonal weighting matrix instead of the optimal weighting matrix because it has better small sample properties (Altonji and Segal, 1996).

4.6 Results and model fit

Estimation results. I report second-stage parameter estimates in Table 3. Time preferences estimates in my baseline specification (column 1) are broadly in line with values estimated in other lifecycle settings. The quarterly discount factor $\delta = 0.987$ (or 0.949 annually) is similar to estimates from Gourinchas and Parker (2002), who find annual values between 0.930 and 0.962. The estimated elasticity of intertemporal substitution $\sigma = 0.435$ is also standard in the literature.[24] The opt-out cost is estimated at $254 every time a worker changes her contribution rate, with a small standard error.

Robustness and sensitivity. My results are robust to using a different weighting matrix or a different set of moment conditions. In column 2 of Table 3, I show that my estimates are very similar when using the inverse of the full variance-covariance matrix of second-stage moment conditions as a weighting matrix. In column 3, I estimate the model using only data on workers hired before the introduction of auto-enrollment. The opt-out cost is less precisely estimated—with a standard error three times larger—because non-participation in the retirement savings plan could either imply a high opt-out cost or a low preference for saving. In contrast, column 4 shows that I obtain parameter estimates that are very similar to the baseline results when I instead use data for autoenrolled workers only. In this case, the bunching of contributions at the 3% default contribution rate can identify the size of the opt-out cost separately from time preferences. In Appendix H, I report and discuss the sensitivity of my estimates using the measure proposed by Andrews et al. (2017).

Model fit. The model fits the data well despite requiring the estimation of only three preference parameters. I report in Figures 3 and 4 the simulated behavior alongside the empirical moments used in estimation. Figure 3 shows that under auto-enrollment, the model generates realistic bunching of contributions at the default contribution rate (3% of salary) and at the match threshold (6% of salary). However, the model tends to over-predict contribution at the match threshold among workers without auto-enrollment. Figure 4 shows that the model replicates the evolution of contributions and participation in the first four years of tenure. Participation increases with tenure for non-autoenrolled workers (Figure 4, left panel) and stays constant at a high rate for autoenrolled workers (Figure 4, middle panel). Over time, autoenrolled workers move away from the default contribution rate and, by the end of the fourth year of tenure, only around 30% of them contribute

[24] Havránek (2015) reports a mean estimate of the elasticity of intertemporal substitution centered around $0.3 - 0.4$, after correcting for publication bias, in a meta-analysis of 169 published studies.

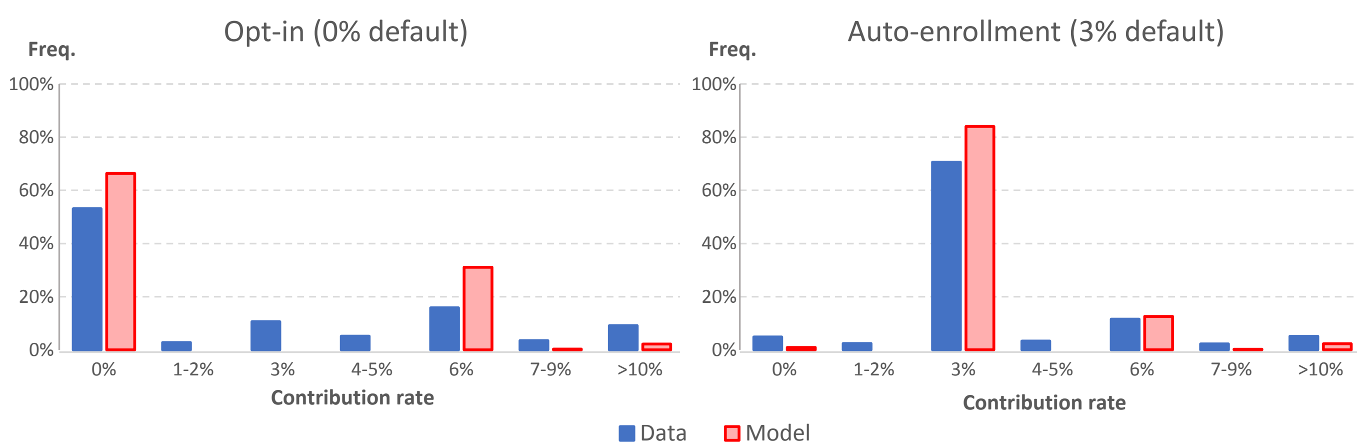
Table 3: Preference parameter estimates

	(1) Baseline model	(2) Full var-cov weighting matrix	(3) Opt-in workers only	(4) Autoenrolled workers only
k	\$ 254	\$ 268	\$ 340	\$ 258
	(11)	(17)	(29)	(11)
δ	0.987	0.987	0.988	0.987
	(0.000)	(0.001)	(0.001)	(0.001)
σ	0.455	0.444	0.454	0.426
	(0.013)	(0.015)	(0.027)	(0.012)
χ^2 stat.	586	583	414	131
(df)	41	41	13	25

Notes: Column 1 presents baseline estimates with a weighting matrix equal to the inverse of the diagonal of the estimated variance-covariance matrix of the second-stage moment conditions (Altonji and Segal, 1996). Column 2 shows results from estimating the model using the inverse of the full variance-covariance matrix of the second-stage moment conditions as a weighting matrix. Column 3 (column 4) presents estimates using only data from workers hired in the 12 months before (after) the introduction of auto-enrollment. The last row shows results from the χ^2 overidentification test with associated degrees of freedom. Standard errors are in parentheses. Data source: administrative 401(k) records from 34 U.S. firms.

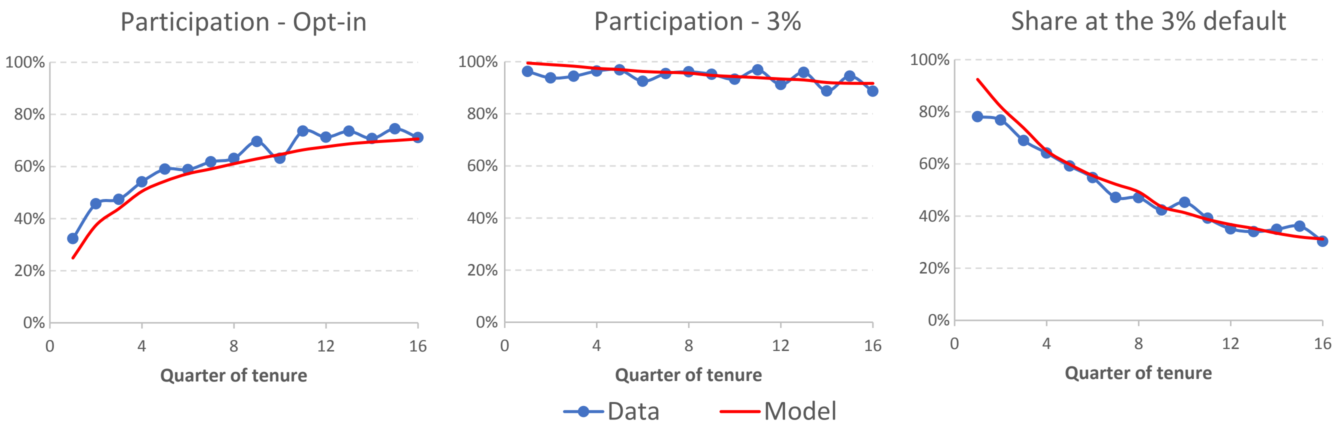
exactly 3% (Figure 4, right panel). While the model is estimated using 401(k) data only, I show in Appendix C.2 that my estimates are also consistent with the evolution of total wealth accumulation over the lifecycle in the Survey of Consumer Finances.

Figure 3: Empirical and simulated distribution of contribution rates in the first 12 months of tenure



Notes: The left (right) panel shows contribution rates of individuals hired in the 12 months before (after) the introduction of auto-enrollment for new hires. The model series corresponds to simulations for identical individuals under a 0% and a 3% default. Data source: administrative 401(k) records from 34 U.S. firms.

Figure 4: Empirical and simulated contribution behavior over 4 years of tenure



Notes: The left (middle) panel shows the 401(k) participation rate of workers hired in the 12 months before (after) the introduction of auto-enrollment for new hires. The right panel shows the share of workers hired under auto-enrollment who contribute exactly 3% of salary (i.e. the auto-enrollment default). The model series are simulations for identical individuals under a 0% and a 3% default. The empirical and simulated sample at each tenure level correspond to workers still employed by the firm at that time. Data source: administrative 401(k) records from 34 U.S. firms.

Overidentification test. The model is formally rejected by the χ^2 overidentification test (reported with associated degrees of freedom in the bottom row of Table 3).[25] The failure of the overidentification test appears to be driven by the fact that: (i) the empirical moments are, in general, estimated precisely with a small variance, and (ii) the model both over-predicts contribution at the match threshold and under-predicts contributions at 3% for workers hired before auto-enrollment.

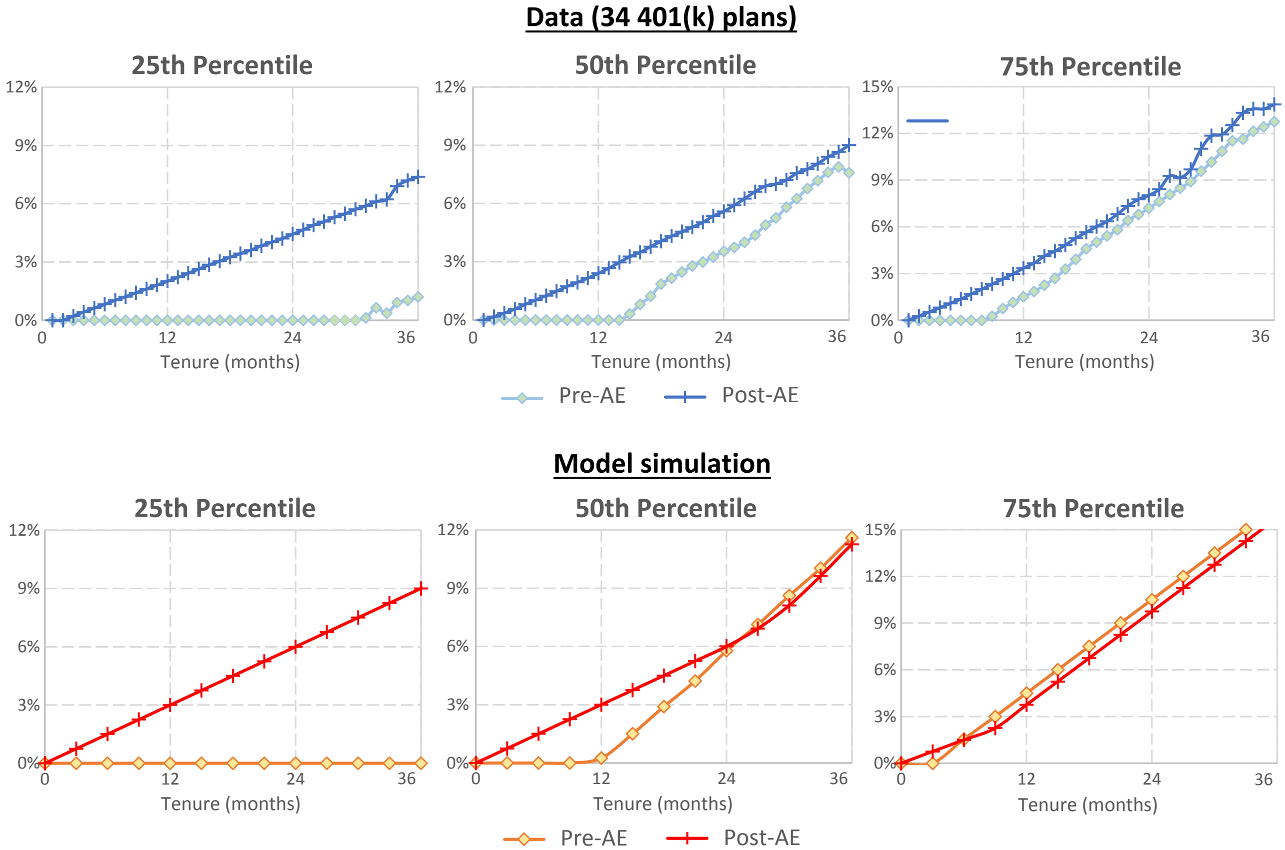
4.7 Heterogeneity in short-run effects

My model is quite parsimonious, with only 3 preference parameters estimated on population averages, but I show that it replicates the heterogeneity in auto-enrollment effects observed in the data. Capturing this heterogeneity is important for making predictions about the policy's long-term effect.

First, I show that, in line with Fact II of Section 2, the auto-enrollment policy's effect on retirement savings is large at the bottom of the distribution of cumulative employee contributions and small at the top. In addition, the median cumulative contributions of non-autoenrolled workers catch up to the median cumulative contributions of autoenrolled workers after three years (Figure 5).

[25] Failure of the overidentification test is common in the estimation of lifecycle models (Gourinchas and Parker, 2002; Laibson et al., 2017).

Figure 5: Empirical and simulated cumulative employee 401(k) contributions to annual salary ratios at the 25th, 50th and 75th percentiles.



Notes: The empirical series are the same as in Figure 1. The model series correspond to simulations for identical individuals under a 0% and a 3% default. The empirical and simulated sample at each tenure level are workers still employed by the firm at that time. Data source: administrative 401(k) records from 34 U.S. firms.

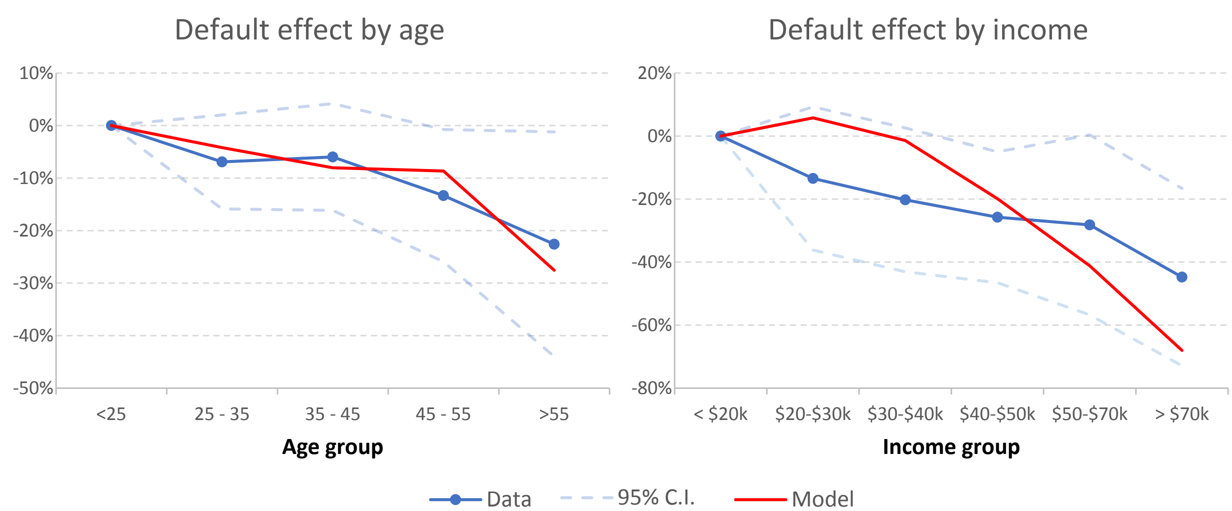
What causes auto-enrollment effects to be different across people? Other studies have suggested that the opt-out cost may be heterogeneous across individuals (as in Bernheim et al., 2015) or that it is stochastic and varies over time (as in Carroll et al., 2009 and DellaVigna, 2018). While these assumptions are realistic, I show in this section that a significant fraction of the observed heterogeneity in behavior can be generated endogeneously in a model with homogeneous preferences. In particular, I show that the estimated model can fit and explain the differences in behavior across age and income groups.

I run a regression to estimate the differences in the propensity to stay at the default option across different age and income groups. The outcome variable is the probability that an employee in her first year of tenure contributes exactly 3% (the auto-enrollment default contribution rate) conditional on participating in the plan. I condition on participation because, in the data, salary is only recorded consistently for employees with a positive contribution rate. Because age and income are correlated, I interact auto-enrollment status (based on date of hire) with both age and income group dummies. I estimate the following linear probability specification controlling for both plan and year fixed effects:

$$Pr\left(s_{i,t}=3\,|\,s_{i,t}>0\right)=\alpha_0+\alpha_1.PostAE_i+\sum_{j=1}^{5}\beta_j.PostAE_i\times age_{i,t}^j+\sum_{k=1}^{6}\gamma_k.PostAE_i\times inc_{i,t}^k+\sum_{j=1}^{5}\tilde{\beta}_k age_{i,t}^j+\sum_{k=1}^{6}\tilde{\gamma}_k inc_{i,t}^k+\varepsilon_{i,t} \tag{8}$$

I run the same specification in the empirical sample and in the simulated sample. Estimates for β_j (the treatment interactions with age groups) and γ_k (the treatment interactions with income groups) are reported in Figure 6. In line with previous evidence (Madrian and Shea, 2001), young and low-income workers have the largest increases in their propensity to contribute at the auto-enrollment default contribution rate. The model captures well this heterogeneity; the coefficient estimates in the simulated sample are all within the 95% confidence interval of the empirical coefficient estimates.

Figure 6: Heterogeneity in default effects by age and income



Notes: Each point corresponds to the coefficients of the interaction terms of auto-enrollment status (based on a worker date of hire) with income and age group dummies, estimated according to equation (8). The dependent variable is a dummy equal to 1 if, conditional on participating, a worker observed in her first year of tenure contributes exactly 3% (the auto-enrollment default contribution rate). The area between the dashed lines corresponds to the 95% confidence intervals of the empirical coefficients. Data source: administrative 401(k) records from 34 U.S. firms.

Age heterogeneity. Conditional on participating, workers in their 20s are 20 percentage points more likely to stay at the default contribution rate than workers in their late 50s and early 60s (Figure 6, left panel). In the model, this heterogeneity reflects the fact that the option value of waiting to contribute decreases with age. A young employee choosing not to save in a particular period has plenty of time to accrue the tax and matching benefits by contributing more later in life. In contrast, there is no option value to waiting in the last period prior to retirement. A worker about to retire permanently forgoes the tax and matching benefits of retirement saving if she does not contribute.

Income heterogeneity. The lowest-income workers are 40 percentage points more likely to remain at the default contribution rate than the highest-paid workers (Figure 6, right panel). In the model, this heterogeneity is explained by the fact that (i) the opt-out cost represents a larger share of earnings for low-income individuals and, to a lesser extent, that (ii) high-income individuals have more to gain by contributing to the retirement savings plan because they face a higher marginal tax rate and expect a lower Social Security replacement rate. In Appendix B.2, I quantify these two channels. In particular, I show that a specification of the model where the opt-out cost is proportional to individuals' earnings does not match the empirical heterogeneity in default effects by income.

5 A Modest Opt-out Cost Explains the Evidence

My baseline opt-out cost estimate with time-consistent preferences of $254 is an order of magnitude smaller than what previous studies have estimated (DellaVigna2009 and 2018; Bernheim et al. 2015). Because the size of the opt-out cost determines the long-term effect of auto-enrollment on wealth and welfare, it is important to understand why my estimate differs from previous work. I conduct a number of simulation exercises to identify which features of my model explain this difference. My finding of a low opt-out cost is driven by the dynamic nature of decisions in the model, the limited liquidity of retirement accounts, and (to a lesser extent) uncertainty about future income. I see my work as complementary of previous approaches: the fully-specified dynamic model could be used to calibrate a reduced-from continuation value in a model like the one of Bernheim et al. (2015).

I fix time preferences at their baseline estimated value ($\delta = 0.987$ and $\sigma = 0.455$), and re-estimate the opt-out cost under seven alternative specifications of the model. This procedure permits to evaluate how different modeling assumptions affect to the size of the estimated opt-out cost. The results are reported in Table 4.

Table 4: Opt-out cost estimates under alternative specifications (holding time preferences fixed).

(1) One year of data	(2) Sticky contributions	(3) Combined (1) + (2)	(4) Perfect insurance	(5) Immediate vesting	(6) Penalty-free withdrawals	(7) Combined (1) to (6)
\$ 226	\$ 308	\$ 712	\$ 465	\$ 344	\$ 609	\$ 3,004

Notes: This table shows estimates of the opt-out cost parameter under different specifications of the model. Other preference parameters are kept constant at their baseline estimated value ($\delta = 0.987$ and $\sigma = 0.455$). Column 1 presents the estimated opt-out cost in the baseline model using only empirical moments from the first year of tenure. Column 2 presents the estimated opt-out cost when workers cannot change their contribution rate after the first period in a job. Column 3 combines the assumptions of columns 1 and 2. Column 4 presents the estimate of the model with no income uncertainty. Column 5 presents the estimate of the model under the assumption than employer match contribution vest imediatly. Column 6 assumes that individuals of all ages can withdraw ressources from the DC account when unemployed without a tax penalty or offset of unemployment benefits. Column 7 combines the assumption of all previous columns. Data source: administrative 401(k) records from 34 U.S. firms.

1. **One year of data.** To make my results comparable with previous studies, I re-estimate the opt-out cost under the baseline specification of the model but only use data from employees' first year of tenure.[26] My estimate of the opt-out cost is barely affected (Table 4, column 1). This suggests that the difference between my estimate and previous work is not driven purely by the fact that I consider a longer time frame. If anything, the estimated opt-out cost is smaller when only using data from the first year of tenure (\$226), instead of data from the first four years of tenure in the baseline (\$254).

2. **Sticky contributions.** In this specification, I assume that employees can only change their contribution rate in the first period in a job.[27] Workers can still contribute more in a future job, but they cannot compensate for not saving early in their tenure by contributing more later in their tenure in the same job. Under this scenario, I estimate an opt-out cost of \$308 when using 4 years of data, and \$712 when using data from the first year of tenure only (Table 4, columns 2 and 3). This last estimate of the opt-out cost, roughly three times larger than in the baseline, highlights the combined effect of focusing only on short-run data and limiting the ability of agents to adjust their contribution in the future in explaining why previous studies estimated implausibly large opt-out costs.

3. **Perfect insurance.** In the next model specification, I assume perfect insurance against

[26]Specifically, I use 30 moment conditions targeting the distribution of contribution rates in the first year of tenure under opt-in and under auto-enrollment (i.e. the share contributing 0%, 1%,...,15% of salary). The excluded categories are contributions at 10% of salary under opt-in and under auto-enrollment.

[27]In practice, I set the contribution opt-out cost equal to infinity after the first period of tenure.

income and employment shocks within each cohort. All individuals of a given age group earn the same income regardless of their earnings productivity draws and employment status. Reducing income uncertainty increases the opt-out cost estimate from \$254 to \$465 (Table 4, column 4). First, reducing income uncertainty lowers individuals' precautionary saving motive and makes DC savings more desirable relative to liquid savings. Second, similar to an S-s model, lower uncertainty reduces the inaction range. Workers have less incentive to wait when the future is less uncertain. Together, these two mechanisms induce workers to make more active decisions when uncertainty is lower and, therefore, a larger opt-out cost is required to generate the observed level of inertia around the default.

4. **Immediate vesting.** In this specificaiton, I assume that employer match contributions vest immediatly (that is, the employee receive the full match from the first quarter of tenure). When employer contributions vest immediatly, employees have less incentive to delay participation and a larger opt-out cost is required to match the data. This effect is however relatively small, and the estimated opt-out cost of \$344 is only moderatly larger than the baseline of \$254 (Table 4 column 5).

5. **Penalty-free withdrawals.** Next, I assume that individuals of all ages can withdraw resources from the DC account when unemployed without a tax penalty or reduction in unemployment benefits.[28] The more liquid the DC account is, the more valuable is the employer matching contribution. Under this specification, a large opt-out cost of around \$600 (Table 4, column 6) is needed to rationalize why non-autoenrolled workers would forgo the matching contribution when they can more easily withdraw and consume resources from the DC account. Indeed, if 401(k) accounts were fully liquid, then contributing less than the matching threshold would be strictly inferior to obtaining the full employer match.[29] Absent a large opt-out cost, workers could contribute to earn the maximum match benefit and increase their consumption at any time by withdrawing from the account without penalty. This is the premise of the back-of-the-envelope calculation in DellaVigna (2009, 2018) that sets the minimum opt-out cost needed to rationalize behavior in a standard model equal to the value of potential employer matching contributions.

[28] I explore the role of the unemployment benefit offset in more details in Appendix B.3.

[29] Choi et al. (2011) study a setting where employees could, almost immediately, increase their consumption by contributing at the matching threshold and withdrawing the contribution right away. In practice, such situations are quite uncommon, as it only concerns employees older than $59^1/2$ with fully vested contributions in 401(k) plans that do not place any restrictions on in-service distributions.

6. **Combined effect.** In the last specification, I combine all the mechanisms discussed previously: I estimate the model using only data from the first year of tenure and assuming sticky contributions, perfect insurance, immediate vesting, and penalty free withdrawals. This specification is closest to the previous literature, and I estimate an equally large opt-out cost of $3,004 (Table 4, column 7). This result highlights how modelling explicitely dynamics, uncertainety, and illiquidity can drive down the size of the opt-out cost needed to match the data by an order of magnitude.

6 Long-run Increases in Wealth are Small Except at the Bottom

6.1 The lifetime impact of a universal auto-enrollment policy

I use the estimated model to predict the effect on lifetime savings of a typical auto-enrollment policy of 3% of salary adopted by all employers. This is the most common default contribution rate in U.S. 401(k) plans (Arnoud et al., 2021). I find that auto-enrollment's effect on lifetime savings is small for most people, but the policy increases wealth accumulation at the bottom of the lifetime earnings distribution. Furthermore, in line with the empirical evidence from Chetty et al. (2014) and Beshears et al. (2021), I find modest crowd-out of liquid savings (or increase in insecured debt) generated by the auto-enrollment policy. I explore a richer set of policy counterfactuals with a default contribution rate of 6% or 10% of income in Appendix C.3.

Model simulation. I simulate the long-term effect of a typical auto-enrollment policy at 3% of income if it were adopted by all employers. I assume that the matching formula is the same in all jobs (i.e. a 50% match for contributions up to 6% of salary, with immediate vesting).[30] I define lifetime earnings as the sum of all labor income and unemployment benefits earned between the ages of 22 and 65. I measure the auto-enrollment effect on wealth as the difference between each simulated individual's savings under the universal auto-enrollment policy and the same simulated individual's savings under an opt-in regime. Figure 7 shows the percent change in total wealth at retirement, defined as the sum of liquid and retirement assets at the age of 65, for each decile of lifetime earnings. Figure 8 shows the change in liquid assets, DC pension assets, and total wealth created by the policy throughout the lifecycle for individuals in different quintiles of lifetime earnings.

Auto-enrollment effect on lifetime savings is small for most. For the majority of people, the auto-enrollment effect on wealth is small: wealth at retirement is changed by less than 2% for the

[30]This is the most common employer matching formula during the 2003-2017 period (Arnoud et al., 2021).

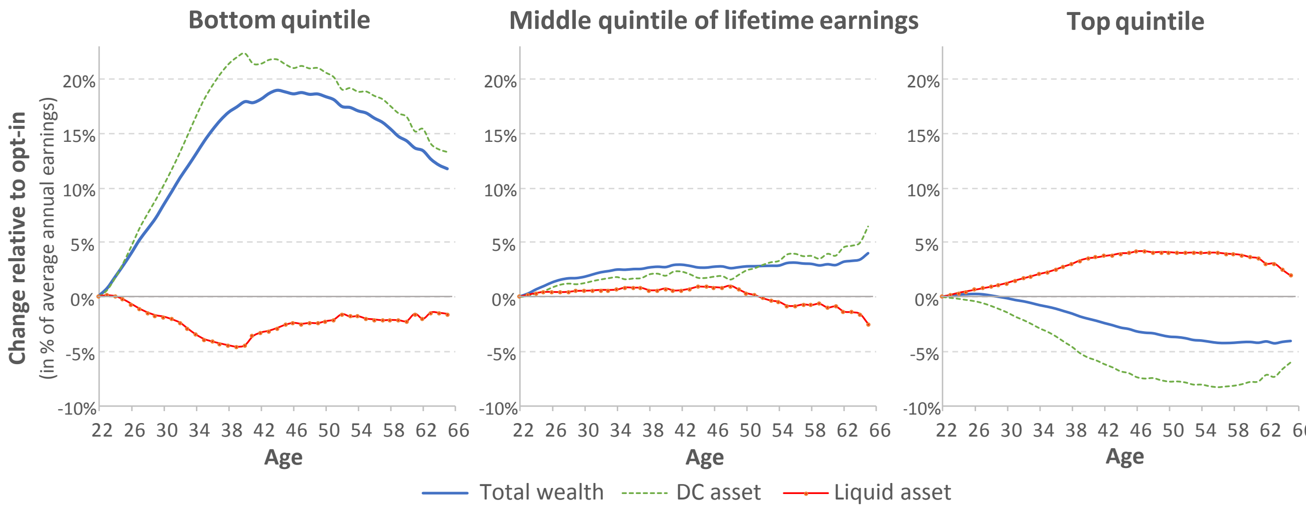
top 7 deciles of lifetime earnings. Furthermore, an auto-enrollment policy at 3% of salary reduces wealth at retirement for individuals in the top two deciles of lifetime earnings (Figure 7). This is because many high-income workers contribute above 3% in the opt-in regime, but stay at the 3% default under auto-enrollment. However, it is important to note that this effect is small: total wealth at age 65 is reduced by less than 1% for individuals in the top two decile of lifetime earnings.

Figure 7: The effect of auto-enrollment on net wealth at age 65

Notes: Each bar corresponds to the model-predicted percentage difference between the sum of retirement and liquid assets at age 65 under an auto-enrollment policy at 3% adopted by all employers versus an opt-in regime adopted by all employers. Changes are expressed as a fraction age 65 wealth in the opt-in regime.

Lifetime savings increase at the bottom. At the bottom decile of the lifetime earnings distribution, auto-enrollment increases total wealth at retirement by more than 12% (Figure 7). Savings increases created by auto-enrollment are concentrated in the first 20 years of working life, and are largest around the age of 45: individuals in the bottom quintile hold an additional 20% of average lifetime earnings under auto-enrollment (Figure 8). In the second half of working life, however, the wealth effect dominates; autoenrolled workers contribute less to the retirement savings account because they have accumulated more wealth early in life. At the bottom of the lifetime earnings distribution, nearly one-third of the increased savings created by auto-enrollment is undone by lower contributions between the ages of 45 and 65. This result suggests that, even with long-term data following individuals over 20 years, empirical estimates of the auto-enrollment effect on wealth may overstate the policy's effect on lifetime savings.

Figure 8: The lifecycle effect of auto-enrollment on retirement savings, liquid assets, and total wealth



Notes: Each series corresponds to the model-predicted difference between assets under an auto-enrollment policy at 3% and an opt-in regime, adopted by all employers. Changes are expressed as a fraction of each simulated individual average annual earnings between the ages of 22 and 65.

Limited crowd-out of liquid savings. As shown in Figure 8, there is limited crowd-out of liquid assets by the auto-enrollment policy. Most increases in retirement savings created by the policy result in increased total wealth accumulation. In the model, the liquid asset (i.e. both liquid wealth and high-interest rate unsecured borrowing) serves a precautionary saving motive, while the DC account serves a lifecycle saving motive. Because the two assets serve different purposes, there is limited substitution between the two. Furthermore, as in Kaplan and Violante (2014), the marginal propensity to consume out of liquid assets is high in the model, so autoenrolled workers reduce their consumption in response to a decrease in disposable income created by the policy. This effect is particularly strong for low-income individuals, who have fewer liquid assets and are more liquidity-constrained. For individuals in the bottom quintile of lifetime earnings, I find that 89% of the increase in retirement savings created by auto-enrollment at the age of 65 passes through to total savings (Figure 8, left panel). In contrast, there is more crowd-out of liquid savings for individuals in the middle quintile of lifetime earnings, as these individuals face less severe liquidity constraints. Still, 62% of the increase in retirement savings created by auto-enrollment at the age of 65 translates into higher total wealth accumulation for individuals in the middle quintile of lifetime earnings (Figure 8 middle panel). These results align with the empirical evidence from Beshears et al. (2021) and Chetty et al. (2014). Beshears et al. (2021) find that auto-enrollment caused no significant increase in unsecured debt after four years. Similarly, Chetty et al. (2014), studying a different policy in Denmark, estimate an 80% pass-through of retirement savings contributions to total savings.

6.2 Discussion of the counterfactual policy analysis

While the model attempts to capture the main aspects of the retirement saving environment, it abstracts from a number of features. In this section, I discuss how the absence of some of these elements could affect the counterfactual policy exercise.

Endogenous labor supply. Employees may be able to undo some of the effect of auto-enrollment by adjusting their future labor supply or shifting the timing of their retirement by a few weeks or months. Introducing these additional margins of adjustment would likely reinforce the main findings of the paper: when employees can adjust both their future contributions and labor supply, the size of the opt-out cost needed to generate the observed inertia at the default and the long-term effect of auto-enrollment on wealth may become even smaller. However, since I am not able to estimate the causal effect of auto-enrollment on labor supply and retirement decisions in the data, I make the more conservative assumption that auto-enrollment has no effect on labor supply and retirement timing.

Default asset allocation. Under auto-enrollment the default contribution rate is often combined with a default asset allocation. In this paper, I study the effect of changing default contribution rate independently from the choice of a default asset allocation. Changing the default asset allocation could have a separate impact on wealth accumulation and welfare that is potentially large but beyond the scope of this paper.[31]

General equilibrium interest rate. A national auto-enrollment policy could raise aggregate saving and lead to a reduction in the equilibrium interest rate. Given the estimated value of the EIS, this general equilibrium channel would likely reduce the effect of auto-enrollment on individual saving. However, given my finding that auto-enrollment has a very modest impact on long-term asset accumulation and that this effect is concentrated at the bottom of the income distribution (a population which accounts for a small share of aggregate savings), this general equilibrium effect is likely to be very small.

[31] Dahlquist et al. (2018) estimate large welfare gains from implementing an optimal default asset allocation in the context of Sweedish pension plans.

7 Auto-enrollment Effects are Stable Across Different Settings

The main advantage of structural estimation is the ability to extrapolate effects to a different policy, population, or institutional environment. In this section, I show that my estimates predict behavior reasonably well in two different contexts. First, I validate my estimates in 86 other U.S. firms. Second, I use estimates from the U.S. sample to predict U.K. workers' response to a national policy rollout. The out-of-sample performance of the model gives credibility to its long-term predictions.

7.1 Out-of-sample validation I: other default contribution rates

The model—estimated using the introduction of auto-enrollment at 3% of salary in 34 U.S. firms—predicts behavior out-of-sample in the 86 U.S. firms used to document Fact I of Section 2. These 86 U.S. firms increased their default contribution rate from a positive default to an even higher default. As shown previously in Section 2.2, worker' response to increasing the default contribution rate is crucial for identifying the mechanism through which defaults affect behavior. I compare the model predicted response of contribution rates to these changes in the auto-enrollment default with the corresponding empirical estimates in Appendix C.1. I cannot reject at the 10% level that the model prediction is equal to the empirical estimate in 8 out of the 11 considered cases (covering 71 out of the 86 firms and more than 85% of workers in the sample).

7.2 Out-of-sample validation II: U.K. national auto-enrollment policy

I exploit the rollout of the U.K. national auto-enrollment policy to validate my estimates in a completely different population and institutional environment (details of the policy rollout are described in Appendix E). I use this policy variation in two ways. First, I compare the contribution choices of workers within the same firms right before and right after the auto-enrollment rollout to the model predictions. Second, I compare the model prediction to the empirical estimate of auto-enrollment's effect after changing jobs (Fact III of Section 2).

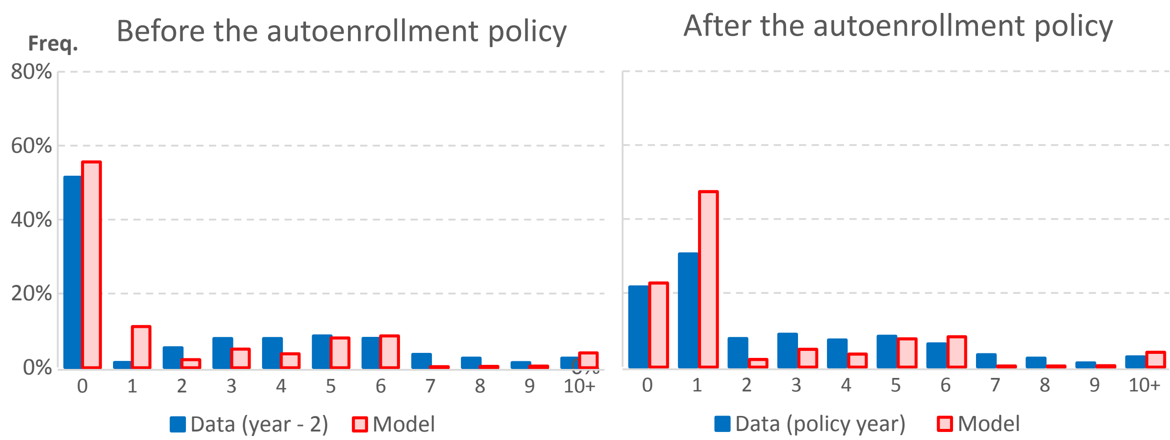
U.K. calibration. I calibrate the model to match the U.K. tax, unemployment insurance, and public pension systems. Of particular interest, the calibration captures the fact that unlike in the U.S. (i) early withdrawals from a retirement savings account are not allowed in the U.K., and (ii) the average replacement rate from public pensions is lower in the U.K. than in the U.S. I use individual-level data from ASHE (prior to the beginning of the auto-enrollment rollout) to estimate the heterogeneity of employers' contribution formulas across firms and the parameters of the income process (with the same two-stage procedure described in Appendix D). I group employers into 5 types based on their employer contribution formula. Other first-stage parameters are set equal to their value in the U.S. calibration. Details on the data and calibration of the U.K. counterfactual are discussed in Appendix E.

Preference parameters. I set time preference parameters equal to their estimated value in the U.S. sample ($\delta = 0.987$ and $\sigma = 0.435$). Using the average exchange rate between the U.S. dollar and the British pound over the sample period (2006 to 2017), I set the opt-out cost $k = £160$.

7.2.1 Within-job evidence

Empirical evidence. I compare workers' participation and contribution behavior before and after the policy rollout. The policy year corresponds to the first survey year in which the policy is binding for a given firm.[32] I compare contributions after the policy implementation to contributions in the same firms two years prior to the policy.[33]

Figure 9: Distribution of employees' contribution rate before and after the U.K. national policy



Notes: The empirical series correspond to workers observed two years before the auto-enrollment mandate became binding at a given firm (left panel) and in the first survey year after the policy (right panel). Data source: U.K. Annual Survey of Hours and Earnings waves 2010 to 2016.

Model exercise. In the model, the policy is implemented at a randomly chosen quarter of employment t^{exp}. Prior to t^{exp}, all employers offer an opt-in regime. After t^{exp}, all employees receive

[32] An alternative approach to this event study is to exploit the staggered implementation of the policy by employer size. Cribb and Emmerson (2016) document very similar patterns in a difference-in-difference specification that takes advantage of this policy variation. I favor a single difference approach because it is easier to compare to model simulations.

[33] Results are identical when using data three, four or five years prior to the policy. I exclude the year prior to the policy because some firms could have adopted auto-enrollment a few months ahead of the legal deadline to do so.

a 1% employer contribution and those with a zero default contribution rate (both new hires and seasoned employees who did not contribute in period $t^{exp} - 1$) see their default contribution rate increased to 1% of income. They can choose to contribute this amount or incur the opt-out cost and change their contribution rate. The policy is unexpected and, after t^{exp}, all individuals expect the new auto-enrollment regime to persist in the future.

Results. The model, with parameters estimated in the U.S., predicts reasonably well workers' contribution behavior in the context of the U.K. national auto-enrollment policy. The model matches the change in the distribution of contribution rates; in particular it generates the fact that both participation and contribution at the 1% default increased by roughly 30 percentage points following the policy implementation (Figure 9).

7.2.2 Across-jobs evidence

Fact III of Section 2 shows that auto-enrollment causes workers to contribute less in their next employer's opt-in retirement savings plan. The model correctly predicts the magnitude of this effect.

Model exercise. For each simulated individual, I change the enrollment regime faced in the first two jobs after a randomly drawn quarter t^{exp}. I assume that all employers have an opt-in enrollment regime prior to t^{exp}. I simulate the model under four scenarios: (i) AE to nonAE: auto-enrollment in job number 1 (in period t^{exp} for both new hires and seasoned non-participating employees) and opt-in in job number 2, (ii) nonAE to nonAE: opt-in in both jobs, (iii) AE to AE: auto-enrollment in both jobs, and (iv) nonAE to AE: opt-in in job number 1 and auto-enrollment in job number 2. Starting from job number 3 after t^{exp}, all employers auto-enroll their new hires. I compute the difference in contributions in the first year of tenure at job number 2 between scenarios (i) and (ii) to obtain the model counterpart to the coefficient AEtononAE in Table 2. I compare scenarios (iii) and (iv) to obtain the model counterpart to the coefficient AEtoAE in Table 2.

Results. The model predicts that auto-enrolling workers in job number 1 causes them to participate 9.6% less in their first year when job number 2 has an opt-in regime. This prediction aligns well with the empirical estimate: participation was 12.8% (*s.e.* 0.055) lower among the previously autoenrolled relative to other new hires in the same opt-in firms (Table 2, panel A). However, the model predicted decline in the contribution rate of 0.11% of salary is smaller than the empirical estimate of a 0.55%(*s.e.* 0.226) of salary (Table 2, panel B). Interestingly, the model prediction for both the participation and contribution rate are very close to the point estimate for new hires in jobs who don't expect be auto-enrolled in the next 12 months: respectively, a drop in participation of 7.5% ($s.e.\,0.052$) and in the contribution rate of 0.12% ($s.e.\,0.266$) (see Appendix Table 13 column 6). For employees moving from an auto-enrollment employer to another auto-enrollment employer (i.e. AE to AE relative to nonAE to AE), the model predicts a modest drop in participation of 3.1% and a small drop in contributions of 0.06% of salary. This is broadly consistent with the empirical finding of no statistically significant difference between previously autoenrolled and non-previously autoenrolled workers when the new employer has auto-enrollment (coefficient AEtoAE in Table 2).

8 Welfare Analysis and Optimal Defaults

In this section, I characterize the optimal enrollment policy under alternative assumptions about social preferences and policy incidence.

8.1 Relationship to the previous literature

A number of recent papers have studied the welfare impact of auto-enrollment policies (Carroll et al., 2009; Bernheim et al., 2015; Goldin and Reck, 2020; Bernheim and Gastell, 2020; Zhong, 2020). This paper adds to this literature by considering: (i) the lifetime welfare impact of auto-enrollment, and (ii) employers' endogenous response to the policy. A central question has been how to identify individuals' normative preferences in a setting where the axiom of revealed preferences often fails. For instance, a worker may contribute at the savings default either because it is equal to her preferred contribution rate or because of inertia. There are two approaches in the literature for assessing welfare in such settings. A first approach, formalized by Bernheim and Rangel (2009), provides bounds on the welfare by inferring normative preferences from a subset of decisions identified as welfare relevant (Bernheim et al., 2015, and Goldin and Reck, 2020). Extending the Bernheim and Rangel framework to a dynamic setting is an interesting avenue for future research, but lies beyond the scope of this paper.[34] I adopt a second approach, similar to Carroll et al. (2009), which consists of building a positive model of behavior that fully specifies the mapping from decision to normative preferences.

[34] A particular challenge is that, in my dynamic setting, the default (i.e. the decision frame in the Bernheim and Rangel framework) is endogenous to individuals' past decisions. After the first period of employment, the default is equal to each employees' contribution choices in the previous period, and workers make their decisions taking into account the fact that their choice today will determine the default they will face in the future.

8.2 Optimal Default Options

I study the optimal selection of a default contribution rate from the perspective of a benevolent social planner.[35] This problem captures the choice governments face when adopting a national auto-enrollment policy (as is the case in the U.K., New Zealand, and Turkey) or the decision U.S. states must make when setting up a state-run auto-IRA program.[36] In the context of 401(k) plans in the U.S., the selection of a default contribution rate is left to employers. Nevertheless, the government can encourage employers to adopt a specific default contribution rate and matching formula by granting qualifying plans an exemption from 401(k) discrimination testing.[37]

8.2.1 Social preferences

The policymaker selects a default contribution rate d^{SP} to be adopted by all employers. I allow social preferences to differ from individual preferences along three dimensions. First, the social discount factor δ^{SP} can be different from individuals' estimated discount factor $\hat{\delta}$. For instance, the policymaker can be more patient than individuals if $\delta^{SP} > \hat{\delta}$. Second, I capture social preferences for redistribution with exogenous Pareto weights φ following Saez (2002). If φ is equal across individuals, the social planner has no redistributive motive beyond declining marginal utility of consumption. Finally, following Goldin and Reck (2020), I assume that the policymaker considers that only a fraction π of the opt-out cost is welfare-relevant. $\pi = 1$ corresponds to the case where the opt-out cost is fully normative (for example, if making a decision involves real costs, such as the fee of a financial adviser) and $\pi = 0$ if the opt-out cost is not welfare-relevant (for instance, if inertia reflects employees' mistakes from the policymaker's perspective).[38]

The policymaker selects a default contribution rate d^{SP} to maximize a social welfare function:

$$\max_{d^{SP} \in S} \left\{ V_0^{SP}\left(d^{SP}\right) \right\}$$

where $V^{SP}\left(d^{SP}\right)$ is the *normative* expected lifetime utility from the perspective of period 0 (i.e.

[35] A different problem, studied by Bubb and Warren (2018), considers the selection of a default contribution rate by a profit-maximizing employer. In their setting, employers exploit workers' mistakes and set an inefficiently low default contribution rate.

[36] To date, seven U.S. states (Oregon, California, Colorado, Connecticut, Illinois, Maryland, and New Jersey) have adopted legislation to extend auto-enrollment to employees with no access to a 401(k) plan through state-run Automatic Individual Retirement Accounts.

[37] Currently, a minimum default contribution rate of 3% is required for a 401(k) plan to become a "Qualified Auto-enrollment Contribution Arrangement" and earn safe harbor status from 401(k) non-discrimination testing (along with other conditions on plan features and employer matching).

[38] Goldin and Reck (2020) discusses the empirical plausibility of different values of π.

before any uncertainty is realized):[39]

$$V^{SP}\left(d^{SP}\right)=\mathbb{E}_0\left[\sum_{t=1}^{A}\left(\delta^{SP}\right)^t\left(\prod_{k=1}^{t-1}\left(1-m_k\right)\right)\varphi_t u_t\left(c_t-\mathbb{1}_{(s_t\neq d_t)}\pi k\right)\right]$$

subject to individuals' optimization defined in Section 3.9.

$$\{c_t, dc_t, l_t\} = \underset{dc_t, l_t}{argmax}\ V_t\left(\vec{X}_t\right)$$

and subject to an aggregate employers' budget constraint:[40]

$$Pf\left(d^{SP}\right)+W\left(d^{SP}\right)+Mtc\left(d^{SP}\right)=\overline{Y} \tag{9}$$

where $Pf\left(d^{SP}\right)$ measures aggregate profits, $W\left(d^{SP}\right)$ aggregate wages and $Mtc\left(d^{SP}\right)$ aggregate employer matching contributions.

Welfare metric. I express changes in welfare relative to the opt-in regime in term of lifetime consumption-equivalent $\gamma(d)$:

$$1+\gamma(d)=\left(\frac{V^{SP}\left(d^{SP}=d\right)}{V^{SP}\left(d^{SP}=0\right)}\right)^{\frac{\sigma}{1-\sigma}}$$

where σ is the elasticity of intertemporal substitution. The policymaker is indifferent between: (i) an auto-enrollment policy with a default d and (ii) an opt-in regime with consumption multiplied by a factor $(1+\gamma(d))$ in every period of life and in every state of the world.

8.2.2 Incidence and means of balancing the budget

I assume that all employers share the same matching formula (initially, a 50% match for contributions up to 6% of income with immediate vesting). Because more workers participate and contribute under auto-enrollment, the policy has the potential to increase employers' matching costs. In order to satisfy employers' aggregate budget constraint (9), I assume that employers respond to the policy in one of three ways:

Reducing profits. The first scenario assumes that the policy's incidence falls completely on employers who reduce their profits to cover higher matching costs. This reduction in profits is

[39] Note that because preferences are homogeneous, lifetime utility is ex-ante identical across individuals.

[40] Note that since employers are ex-ante identical in this set-up, the aggregate employers' budget constraint is equal, in expectations, to each individual employer's budget constraint.

assumed to have no direct effect on social welfare and no effect on the interest rate (which is calibrated to match the long-term yield on government bonds). This assumption is implicit in previous characterizations of an optimal default contribution rate that did not feature employers' budget constraints (Bernheim et al., 2015; Goldin and Reck, 2020). A drawback of this specification is that the direct effect of auto-enrollment on welfare cannot be disentangled from the effect of employers increasing workers' total compensation by spending more on matching.

An alternative scenario assumes that the incidence of the policy falls fully on workers. I explore two possible specifications:

Reducing the match rate. Butrica and Karamcheva (2012) document that, in the cross-section, auto-enrollment is correlated with a less generous matching formula and that total retirement plan costs and total compensation are not different between firms with and without auto-enrollment. Motivated by this evidence, I assume that employers adjust the match rate $match^e$ to satisfy the aggregate budget constraint (9). The match rate adjusts to keep employers' total matching expenditure constant under alternative default contribution rates.

Reducing wages. Another possibility is that employers reduce wages to offset higher matching costs. I assume that all wages are reduced by the same proportional factor in order to satisfy employers' aggregate budget constraint (9). Employers adjust the wage level to keep the total compensation bill $W\left(d^{SP}\right) + Mtc\left(d^{SP}\right)$ constant.

8.3 Optimal Policies under Alternative Specifications of Social Preferences

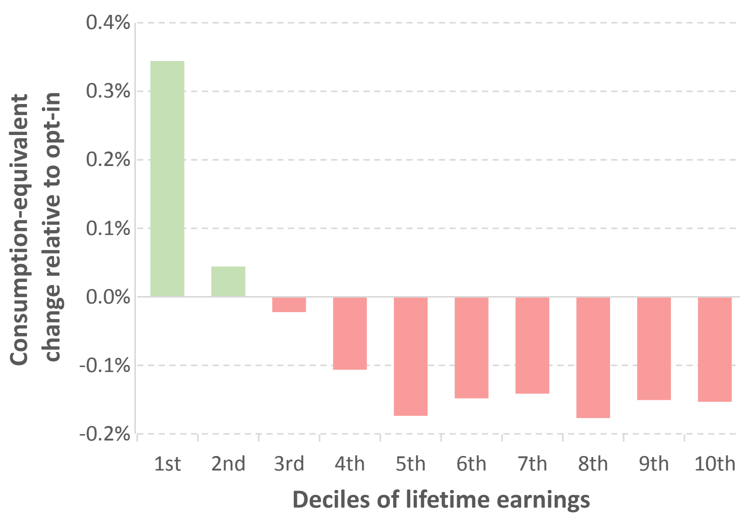
I solve for the optimal default contribution rate under three specifications of social preferences: a utilitarian, inequality-averse, and paternalistic policymaker. In each case, I allow for the opt-out cost to be either fully normative (i.e. $\pi = 1$) or welfare irrelevant (i.e. $\pi = 0$). Optimal policies, as summarized in table 5, are identical under both values of π.

8.3.1 A utilitarian policymaker prefers the opt-in regime to auto-enrollment

I consider first the case of a utilitarian policymaker who shares the same discount rate as individuals ($\delta^{SP} = \hat{\delta}$) and does not have redistributive preferences beyond declining marginal utility ($\varphi_t = 1$). The utilitarian policymaker always prefers an opt-in regime to auto-enrollment, even when the policy incidence is fully on employers (Table 5). The matching and tax incentives for retirement saving cause individuals to consume too much in retirement relative to the ideal consumption plan implied by their time preference parameters. Hence, auto-enrollment, by inducing people to save even more for retirement, tends to reduce utilitarian welfare (Appendix Figure 17 panel A). This negative auto-enrollment effect on welfare is largest if employers reduce wages to compensate for higher matching contribution costs because, from the perspective of utilitarian welfare, wages are always preferred to matching contributions (Appendix Figure 17 panel C).[41] If the opt-out cost is not welfare-relevant ($\pi = 0$), an auto-enrollment default at the maximum contribution rate (15%) is almost as good as the opt-in regime (Appendix Figure 17 panel A). The intuition for this result is that a very high default induces more workers to incur the opt-out cost and make an active decision (what Carroll et al. (2009) refer to as an "offset default").

The average auto-enrollment effect on welfare masks important heterogeneity across the income distribution. As shown in Figure 10, even absent any paternalistic motive, the policy has a significant positive effect at the bottom: a 6% auto-enrollment default increases welfare by 0.3% in lifetime consumption-equivalent at the bottom of the lifetime income distribution when the policy incidence is fully on employers and the opt-out cost is fully welfare-relevant. Because the opt-out cost represents a larger share of earnings for low-income individuals, they are more likely to forgo the benefits of retirement saving absent auto-enrollment.

Figure 10: Heterogeneity in auto-enrollment's effect on lifetime welfare - Utilitarian policymaker



Notes: Each bar corresponds to the difference between lifetime utility under an auto-enrollment policy at 6% adopted by all employers and an opt-in regime for individuals in each decile of lifetime earnings. Changes are expressed in lifetime consumption-equivalent. The policy incidence is fully on employers, and social preferences are assumed to be equal to individuals' preferences ($\delta^{SP} = \hat{\delta}$ and $\varphi_t = 1$) with a fully welfare-relevant opt-out cost ($\pi = 1$).

[41] The reduction in utilitarian welfare from adopting a 6% default is nearly 50% larger if employers reduce wages instead of lowering the match rate.

8.3.2 An inequality-averse policymaker prefers auto-enrollment

In what follows, I show that a policymaker who puts more weight on low-income individuals will select a default contribution rate near the employer matching threshold even if the policy incidence falls fully on workers.

Pareto weights. I adopt a functional form similar to Saez (2002) with Pareto weights equal to

$$\varphi_t = \bar{y}\left(\theta_t, a_t, emp_t\right)^{-\nu}$$

where $\bar{y}$ is the income under the opt-in regime of an individual at age a_t, with labor productivity θ_t, and employment status emp_t. If $\nu = 0$ the social planner has no redistributive motive beyond declining marginal utility of consumption, while $\nu > 0$ captures the degree of inequality aversion. These Pareto weights are treated as exogenous and do not vary across model counterfactuals. Following Saez (2002), I set $\nu = 1$. This specification creates a motive for both intra-generational redistribution (from high to low labor productivity types θ) and inter-generational redistribution (from the relatively rich middle aged to the relatively poor young and elderly).

Optimal policy. When the incidence of the policy falls on employers (through a reduction in profits) the optimal default selected by an inequality averse policymaker is equal to the employer matching threshold (6% of income). However, when the incidence of the policy falls on workers, the policymaker must trade off (i) improving the welfare of low-income workers with a higher auto-enrollment default and (ii) reducing the resources available to higher income individuals. A default contribution rate of 5% of income balances these two motives when the policymaker is inequality averse (Appendix Figure 18).

8.3.3 A paternalistic policymaker sets the default at the employer matching threshold

Many countries have instituted forced savings programs (such as Social Security contributions in the U.S) and tax incentives for private retirement saving contributions. The existence of these programs and incentives indicate that policymakers may value retirement consumption more than implied by individuals' revealed preferences. To capture this idea, I consider the case of a paternalistic policymaker who is more patient than individuals ($\delta^{SP} = 1$) but does not have a redistributive motive beyond declining marginal utility ($\varphi_t = 1$). The paternalistic policymaker selects a default contribution rate equal to the threshold on employer matching.

There are multiple reasons why the policymaker may discount the future less than individuals.

First, if the policymaker cannot commit to not helping the elderly living in poverty, individuals may rationally undersave for retirement. In such setting, known in the literature as the Samaritan's dilemma, the optimal policy solves the problem of a committed planner who discounts the future less than the agents (Sleet and Yeltekin, 2006). Second, if people suffer from behavioral biases or have time-inconsistent preferences, a paternalistic policymaker can correct for individuals' mistakes with more future-oriented policies (Laibson, 1997).[42] Finally, a lower social discount rate may be desirable on normative grounds. This argument goes back to Ramsey (1928), who contended that discounting the future was "*ethically indefensible*". Similarly, Caplin and Leahy (2004) argue that the normative appeal of the principle of revealed preferences may not extend to dynamic settings and that, in most cases, policymakers should be more patient than individuals.

Optimal defaults. An auto-enrollment default at the match threshold (6% of income) is always preferred by the paternalistic policymaker (Appendix Figure 19). Because the paternalistic policymaker values retirement savings more than individuals, he selects the default which maximizes retirement wealth accumulation. A default above the matching threshold reduces asset accumulation: contributions above the threshold are viewed as undesirable by many workers (as they do not earn a matching benefit) and induce a larger fraction of opt-outs. In contrast to the utilitarian case, reducing wages to balance employers' budget constraint is (weakly) preferred to reducing the match rate. This difference arises because, from the perspective of the paternalistic policymaker, allocating more funds to matching incentives and fewer funds to wages can improve welfare by shifting consumption toward the future.

Table 5: Summary table of the optimal default under alternative specifications

		Policy incidence		
		Employers profits	Matching formula	Wages adjustment
Social preferences	Utilitarian	Opt-in	Opt-in	Opt-in
	Inequality averse	AE 6%	AE 5%	AE 5%
	Paternalistic	AE 6%	AE 6%	AE 6%

Notes: Each cell corresponds to the optimal policy under an alternative assumption about social preferences and policy incidence. Opt-in corresponds to a zero default contribution rate $d^{SP} = 0$, and AE corresponds to a positive default contribution rate adopted by all employers over a lifetime. Optimal policies are identical whether the opt-out cost is fully normative ($\pi = 1$) or welfare-irrelevant ($\pi = 0$).

[42]In addition to the current analysis, I characterize the optimal policy when individuals undersave due to present bias in Appendix B.1.2.

9 Conclusion

A growing literature has documented, in a variety of settings, that individuals stick with the default option and make infrequent switching decisions.[43] This paper has shown, in the context of contributions to retirement savings plans, the importance of studying decision-making inertia in a fully dynamic setting. When individuals can compensate for their inaction with other actions in the future, small opt-out costs generate large inertia and the observed short-run effects may overstate the long-term consequences of inaction. These insights are likely to extend, perhaps to different degrees, to other settings where large opt-out costs have been estimated. In similar settings, my results suggest that freedom of choice is preserved not only because individuals can offset the effect of a nudge by opting out right away, but also because they can take corrective actions in the future. Therefore, the appeal of what Thaler and Sunstein (2003) call "libertarian paternalism" is particularly strong. It is instructive to contrast my setting with other domains, such as healthcare plan or school choice, in which future actions cannot easily compensate for present inertia. In these settings, opt-out costs estimates are likely to remain large. Because there is little difference between a nudge and a hard mandate when opt-out costs are very large, the appeal of libertarian paternalism is limited in these other settings and more traditional policy tools—such as prohibitions and mandates—may be preferable.

[43] For instance, in the retirement savings context, significant inertia and large opt-out costs have been estimated for switching between retirement funds (Illanes, 2016; Dahlquist et al., 2018), or between a defined benefit and defined contribution plan (Dobrescu et al., 2018).

A Parameters Definition

Table 6 summarizes all the model parameters introduced in the paper (excluding the appendices).

Table 6: Summary table of the model's parameters

	Preference Parameters		**Assets**
δ	Discount factor	R^{dc}	Rate of return on DC wealth
σ	Elasticity of intertemporal substitution	$R^{liq}()$	Rate of return on liquid assets
k	Opt-out cost	$1+r$	Interest rate on positive asset holdings
	State variables	$1+r^{cc}$	Interest rate on unsecured debt
X_t	Vector of all state variables	$\underline{l}_a$	Borrowing limit
a	Age		**Defined Contribution Account**
emp_t	Employment status	$E()$	Distribution of DC account types
ten_t	Tenure	$ec^e()$	Employer DC contribution
e	Employer DC plan type	$fixed^e$	Fixed employer contribution
θ	Labor productivity	$match^e$	Employer matching rate
ae	Average lifetime earnings	cap^e	Threshold on employer matching
l	Liquid assets	$\Upsilon_e()$	Vesting risk-adjustment
dc	DC wealth stock		**Labor Market**
d	Default contribution rate	$\pi^{JJ}()$	Job-to-job transition probability
$F()$	Joint distribution of θ, emp, e	$\pi^{EU}()$	Unemployment transition probability
	Choices	$\pi^{UE}()$	Out-of-unemployment transition probability
c	Consumption	$\{\delta_i\}_{i=0}^{3}$	Deterministic component of earnings
s	DC contribution rate	w	Labor earnings
$draw$	DC withdrawal rate	ρ	Autocorrelation in earnings shocks
	Utility	ξ	Earnings innovation if continuously employed
$u_a()$	Utility function	$\sigma^2_{\xi_1}$	Variance of the first earnings innovation
$V()$	Value function	σ^2_{ξ}	Variance of subsequent innovations
$V^S()$	Value function cdt. on DC contrib. s	ζ	Earnings innovation after job-to-job transition
	Demographics	μ^{JJ}	Avg. wage gain after a job-to-job transition
A	Maximum age	ς_t	Earnings innovation out of unemployment
A^R	Retirement age	μ^{UE}	Avg. wage loss out of unemployment
m_a	Mortality risk	ι	Measurement error in earnings
n_a	Equivalence scale	σ^2_{ι}	Variance of measurement error
	Welfare analysis	η	Earnings innovation plus measurement error
$V^{SP}()$	Social welfare function		**Tax and benefit system**
$\gamma()$	Consumption-equivalent variation	$tax^i()$	Tax on income
δ^{SP}	Social discount factor	τ^k	Tax on capital returns
π	Degree to which k is welfare-relevant	$limit_a$	Tax limit on DC contributions
φ_t	Pareto weights	pen_a	Tax penalty of early DC withdrawals
ν	Degree of inequality aversion	$ui()$	Unemployment insurance benefit
$\bar{y}()$	Avg. individual income under opt-in	$ss()$	Public pension income
$\overline{Y}$	Employers' aggregate budget constraint		
$Pf()$	Aggregate employers' profits		
$W()$	Aggregate wages		
$Mtc()$	Agg. employers' matching contributions		

APPENDIX FOR ONLINE PUBLICATION

B Model Extensions

B.1 Present-biased preferences

Several empirical studies have documented that measures of time-inconsistency correlate with the propensity to stay at the default option in retirement savings plans (Blumenstock et al., 2018; Brown et al., 2016; Brown and Previtero, 2016, Goda et al., 2020). Under present-biased preferences workers procrastinate on changing their contribution rate because the opt-out cost is paid in the present while the benefits of retirement saving are in the future (O'Donoghue and Rabin, 1998; Carroll et al., 2009; DellaVigna 2009 and 2018). I compare two different specifications of present-biased preferences and do not find support for one in which present bias directly affects contributions.

A first view is that contribution choices only reflect the regular discount factor δ because contribution changes are implemented by employers with a delay—usually at the beginning of the next pay period. As noted by Carroll et al. (2009), this implies that a contribution rate choice serves as a commitment to save in the future, starting from the next paycheck. Under this view, present bias is similar to a higher opt-out cost: it reduces the probability of switching but does not affect the contribution choice conditional on making an active decision. In the absence of a source of variation in my data that would allow me to separately identify this form of present bias from opt-out costs, I interpret my baseline estimates as capturing the combined effect of present bias and other forms of opt-out costs.[44]

A second view is that present bias directly affects contribution preferences. Workers choose low contribution rates because they do not want to reduce their consumption in the present and believe (mistakenly) that they will contribute more later. It is important to explore this possibility because, unlike the first view, the welfare implications may be substantially different from the baseline model. Under this specification of present bias, contribution decisions do not reveal workers long-term preferences and are systematically too low. This creates a motive for the social planner to nudge workers into choosing a higher contribution rate. I explore the role of present bias in the model by introducing an additional discount factor $\beta \leq 1$ between the present and all future periods. I assume that workers are naive about their future selves' present bias. The objective function for a present biased worker V^{PB} is given by equation (10) where V_{t+1} is the value function in the baseline

[44]One potential way to to identify present bias separately from other forms of contribution switching costs could be to use data on both 401(k) contributions and liquid asset holdings. Beshears et al. (2021) study a setting with such linked data.

exponential discounting case.

$$V_t^{PB}(X_t) = \max_{s_t \in S, l_{t+1}} u_a\left(c_t - \mathbb{1}_{(s_t \neq d_t)}.k\right) + \beta\delta(1 - m_a) \int V_{t+1}(X_{t+1})\, dF(\theta_t, emp_t, e_t) \tag{10}$$

B.1.1 Estimation with present biased preferences

I consider two values of the short-term discount factor: (i) $\beta = 0.504$ which is the value estimated by Laibson et al. (2017) in a lifecycle setting, and (ii) $\beta = 0.80$ which is in the high-end of short-term discount factor estimates. I re-estimate the model's three other preference parameters following the same procedure outlined in section 4. Results are reported in Table 7.

Table 7: Preference parameter estimates- Model extension

	(1) Baseline model $\beta = 1$	(2) Present bias I $\beta = 0.800$	(3) Present bias II $\beta = 0.504$
k	\$ 254	\$ 269	\$ 430
	(11)	(9)	(13)
δ	0.987	0.989	0.999
	(0.000)	(0.000)	(0.000)
σ	0.455	0.454	0.625
	(0.013)	(0.014)	(0.017)
χ^2 stat.	586	607	400
(df)	41	41	41

Notes: Column 1 reproduce baseline estimates from Table 3 Column 1. Column 2 presents parameters estimates with a low degree of present bias $\beta = 0.8$. Column 3 presents parameters estimated with a higher degree of present bias $\beta = 0.504$, which corresponds to the value estimated by Laibson et al. (2017). In all columns the weighting matrix is the inverse of the diagonal of the estimated variance-covariance matrix of the second-stage moment conditions. The last row shows results from the χ^2 overidentification test with associated degrees of freedom. Standard errors are in parentheses. Data source: administrative 401(k) records from 34 U.S. firms.

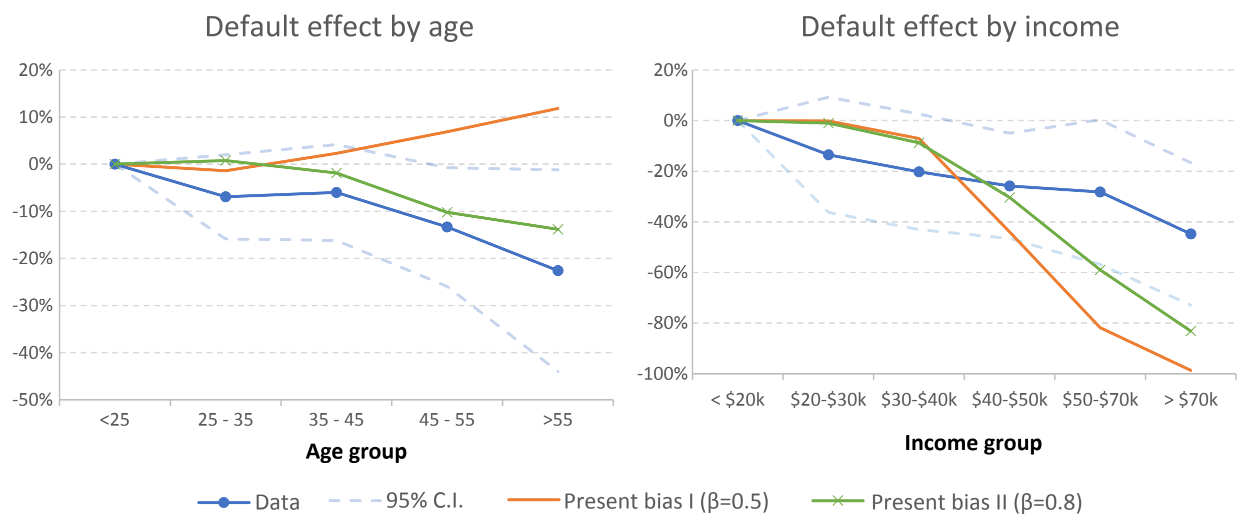
Surprisingly, introducing this form of present-biased preferences leads to a higher estimate of the opt-out cost: \$ 430 for $\beta = 0.504$ and \$ 269 for $\beta = 0.80$. To match the fact that workers contribute to the plan despite their present bias requires a higher long-term discount factor.[45] While adding a short-term discount factor β implies that workers value less the benefits of retirement savings, a higher estimate of the long-term discount factor δ implies that they value these benefits more. The second channel dominates and, because many workers value the retirement period more under

[45] I estimate $\delta = 0.999$ (or 0.996 annually) for $\beta = 0.504$ and $\delta = 0.989$ (or 0.957 annually) which is higher than my baseline estimate of $\delta = 0.987$ (or 0.949 annually)

this parametrization of the model, the opt-out cost needed to rationalize the observed inertia at the savings default is higher. To build intuition for this result, consider the case of a 25-year-old worker who retires at age 65 (after 160 quarters). She discounts utility at retirement by a factor $\beta(\delta)^{160}$. While a lower value of β implies that she values the benefits of retirement saving (including the employer match) less relative to the cost of opting out in the present, a small increase in the value of δ can reverse this result.[46]

In addition, a higher long-term discount factor causes the model not to match well the observed heterogeneity in default behavior by age. When the estimate of δ is close to unity, older worker are more likely to stay at the default option than observed in the data (Figure 11, left panel).

Figure 11: Heterogeneity in default effects by age and income - Present-biased preferences



Notes: Each point corresponds to the coefficients of the interaction terms of auto-enrollment status (based on a worker date of hire) with income and age group dummies, estimated according to equation (8). The dependent variable is a dummy equal to 1 if, conditional on participating, a worker observed in her first year of tenure contributes exactly 3% (the auto-enrollment default contribution rate). The area between the dashed lines corresponds to the 95% confidence intervals of the empirical coefficients. Data source: administrative 401(k) records from 34 U.S. firms.

B.1.2 Weflare with present biased preferences

Next, I solve for the optimal default in the version of the model estimated with naive present-biased preferences. I compute welfare under the assumption that only long-term preferences are normatively relevant as in Carroll et al. (2009). While the long-term criterion is somewhat controversial (Bernheim and Rangel, 2009), it captures the desire for a paternalistic policymaker to counteract

[46] My estimates imply that $\beta(\delta)^{160}$ is lower in the baseline model than in either of the two parametrizations of present bias.

individuals' present biased actions. The optimal default under alternative assumptions about social preferences and about the policy incidence are reported in Table 8. With a moderate level of present bias ($\beta = 0.8$) optimal policies are very similar to those in the baseline model. When individuals are assumed to be more present biased ($\beta = 0.5$) and the opt-out cost is larger ($k = \$430$), the optimal policy is to set a high default contribution rate (around 10% of salary).

Table 8: Summary table of the optimal default under present-biased preferences

		Present-bias I ($\beta = 0.800$)			Present-bias II ($\beta = 0.500$)		
		Employers profits	Matching rate	Wages adjustment	Employers profits	Matching rate	Wages adjustment
Long-term criterion	$\pi = 1$	Opt-in	Opt-in	Opt-in	AE 9%	AE 9%	AE 9%
	$\pi = 0$	AE 15%	Opt-in	Opt-in	AE 10%	AE 10%	AE 10%
Long-term criterion	$\pi = 1$	AE 6%	AE 5%	AE 5%	AE 10%	AE 10%	AE 10%
+ inequality averse	$\pi = 0$	AE 6%	AE 5%	AE 6%	AE 11%	AE 10%	AE 11%

Notes: Each cell corresponds to the optimal policy under an alternative assumption about social preferences and policy incidence for two alternative calibration of the present bias parameter. Opt-in corresponds to a zero default contribution rate $d^{SP} = 0$, and AE corresponds to a positive default contribution rate adopted by all employers over a lifetime.

B.2 Proportional Opt-Out Costs

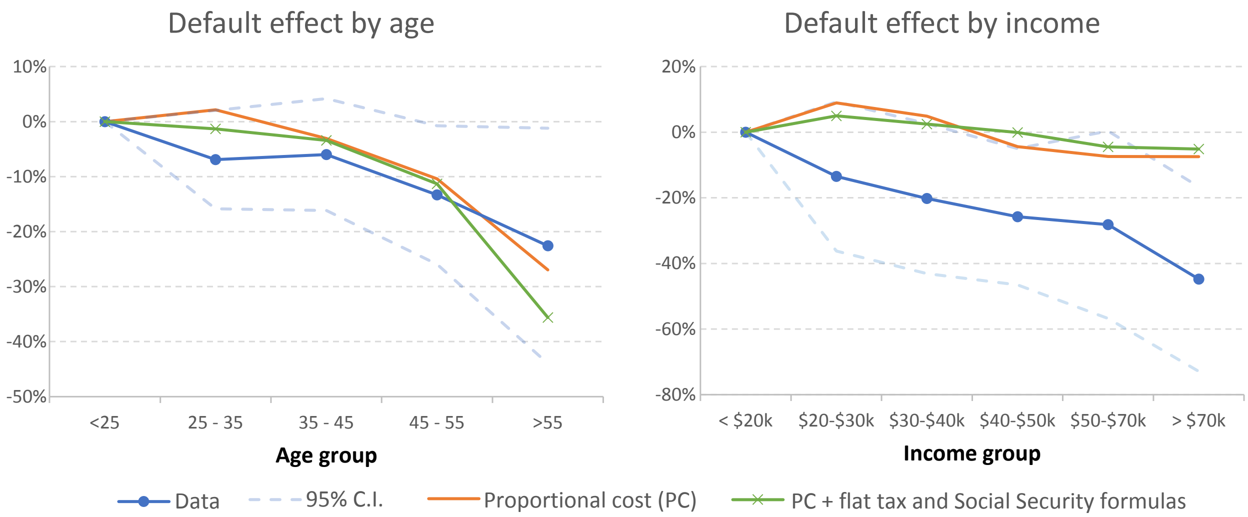
As an extension to the baseline model, I consider the case where the opt-out cost is proportional to each workers' labor earnings. For instance, if changing the DC contribution rate takes time, this specification captures the fact that the opportunity cost of time is higher for highly-paid workers. I introduce an opt-out cost $\tilde{k}$ that is proportional to earnings:

$$u_a \left(c_t - \mathbb{1}_{(s_t \neq d_t)} \tilde{k}.w_t \right)$$

I fix time preferences to their estimated value ($\delta = 0.987$ and $\sigma = 0.455$) and estimate $\tilde{k}$ to be equal to 3.16% of quarterly income (i.e. equivalent to \$ 292 at the average earnings level). Under this specification, high income workers are more likely to stay at the default because the opt-out cost is now larger for them (Figure 12, right panel). Under this parametrization, the model fails to match the empirical heterogeneity in treatment effects across income groups. The relationship between income and the propensity to stay at the default is even flatter under the additional assumption that Social

Security benefits and income tax rates are constant (i.e. I set the Social Security replacement equal to 50% for everyone and the income tax rate at 15%). Results are reported in Figure 12.

Figure 12: Heterogeneity in default effects by age and income - Proportional opt-out cost



Notes: Each point corresponds to the coefficients of the interaction terms of auto-enrollment status (based on a worker date of hire) with income and age group dummies, estimated according to equation (8). The dependent variable is a dummy equal to 1 if, conditional on participating, a worker observed in her first year of tenure contributes exactly 3% (the auto-enrollment default contribution rate). The area between the dashed lines corresponds to the 95% confidence intervals of the empirical coefficients. Data source: administrative 401(k) records from 34 U.S. firms.

B.3 No Offset of Unemployment Benefits

In the baseline specification of the model, early withdrawals from the DC account reduce unemployment benefits. This modeling assumption is consistent with the Unemployment Compensation Amendments of 1976 which requires that all retirement income be offset against unemployment compensation. However there are differences across states in how this offset is implemented. In general, for withdrawals from a 401(k) plan, only the amount contributed by the employer offsets unemployment benefits. For instance, in New Jersey 50% of retirement income, including *"benefits paid in a lump sum such as 401(k)"*, is substracted from the unemployment benefit if both the claimant and the base-period employer contributed to the pension plan.[47]

However the rules are complex and vary across states (see Franco (2004) for a discussion of differences across U.S. states). Depending on the specific way the offset is implemented by states' unemployment insurance agencies, unemployed individuals may be able to avoid this offset by first rolling over their 401(k) assets into an Individual Retirement Account before withdrawing those

[47] Source: https://nj.gov/labor/ui/content/faq.html

resources. As a robustness check, I re-estimate the model under the assumption that 401(k) withdrawals do not offset unemployment insurance benefits. The estimated parameters, in table 9, are little affected by this alternative assumption: the contribution opt-out cost and EIS are slighly higher (\$313 and 0.513 relative to baseline values of, respectively, \$254 and 0.455), and the quarterly discount factor is slighly lower (0.983 relative to an estimate of 0.987 in the baseline model).

Table 9: Preference parameter estimates with no offset of unemployment benefits

(1) k	(2) δ	(3) σ	(4) χ^2 stat.
\$ 313	0.983	\$ 0.513	710
(8)	(0.000)	(0.013)	41 $d.f.$

Notes: Column 1 to 3 presents second stage parameter estimates for the opt-out cost, the quarterly discount factor, and the elasticity of intertemporal subsitution under the assumption of no unemployment offset of early 401(k) withdrawals. The weighting matrix is equal to the inverse of the diagonal of the estimated variance-covariance matrix of the second-stage moment conditions. Column 4 shows results from the χ^2 overidentification test with associated degrees of freedom. Standard errors are in parentheses. Data source: administrative 401(k) records from 34 U.S. firms.

C Additional Model Counterfactuals

C.1 Out-of-Sample Validation: Raising the Default Contribution Rate

In column 1 of Table 10, I disaggregate the evidence presented in Table 1. I estimate 11 separate linear probability regressions for each observed pair of initial and post-increase default contribution rates. The outcome variable is a dummy equal to 1 if a worker contributes strictly below the initial default contribution rate. For instance, if the default is increased from 3% to 6% of salary, this corresponds to contributions at 0, 1%, or 2% of salary.

In column 2 of Table 10, I present the model prediction for each of these changes in the default contribution rate. I use the same calibration and simulation procedure used for estimation in Section 4, but impose different auto-enrollment defaults. I test the difference between the empirical estimates and the model out-of-sample predictions, and report the corresponding p-values in column 5. I cannot reject at the 10% level that the model prediction is equal to the empirical estimate in 8 out of the 11 considered cases (covering 71 out of the 86 firms and more than 85% of workers in the sample).

Table 10: Out-of-sample validation of the model predictions in 86 U.S. 401(k) plans

	Contribution rate < initial default		**Sample size**		**Model vs data**
	(1)	(2)	(3)	(4)	(5)
	Data 86 plans	Model prediction	Nbr. of plans	Nbr. of workers	P-value of difference
Default increased by 1%					
Default 2% → 3%	0.017 (0.014)	0.007	11	31,364	**[0.483]**
Default 3% → 4%	0.016 (0.013)	0.005	10	13,116	**[0.430]**
Default 4% → 5%	-0.003 (0.020)	0.013	3	1,821	**[0.513]**
Default 5% → 6%	-0.016 (0.009)	0.034	5	3,970	[0.005]
Default increased by 2%					
Default 1% → 3%	0.023 (0.025)	0.020	1	1,067	**[0.917]**
Default 2% → 4%	-0.005 (0.011)	0.012	4	1,793	**[0.231]**
Default 3% → 5%	0.022*** (0.005)	0.018	14	56,011	**[0.456]**
Default 4% → 6%	0.031*** (0.007)	0.047	9	17,989	[0.048]
Default 6% → 8%	0.067*** (0.021)	0.148	1	673	[0.000]
Default increased by 3 or 4%					
Default 3% → 6%	0.045*** (0.016)	0.052	26	27,190	**[0.648]**
Default 3% → 7%	0.060 (0.017)	0.132	2	4,219	**[0.146]**
Individual's characteristics	✓				
Plan FE	✓				

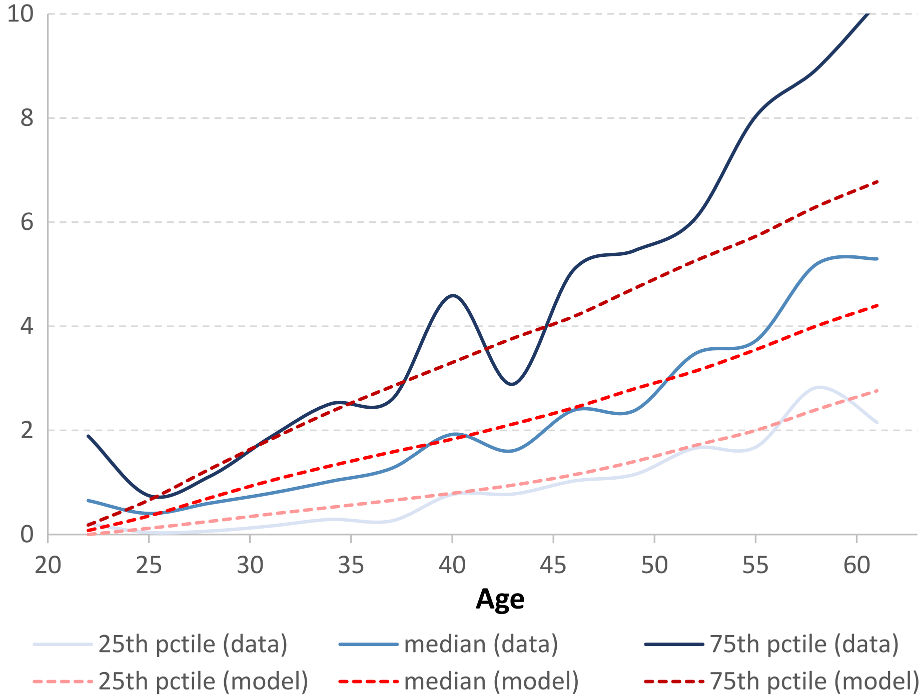
* p<0.10, ** p<0.05, *** p<0.01

Notes: Column 1 reports coefficients from a linear probability regression where the dependent variable is dummy indicating that a worker, observed in her first year of tenure and after the end of the grace period, contributes strictly below the initial default contribution rate. The coefficients correspond to the difference between employees hired before and after each combination of initial and post-increase auto-enrollment default contribution rates. Additional controls for individuals' characteristics include log of tenure, age, age square, and age cube. Column 2 reports the model prediction for each change in the default contribution rate. Columns 3 and 4 report the number of observations in each pair of initial and post-increase default contribution rates. Column 5 reports the p-value of the difference between the empirical coefficient and the model prediction. Standard errors clustered at the firm level are in parentheses. Data source: administrative 401(k) records from 86 U.S. firms.

C.2 Wealth Accumulation over the Lifecycle

While the model is estimated using 401(k) data only, my baseline estimates are also consistent with the evolution of total wealth accumulation over the lifecycle in the 2016 Survey of Consumer Finances (SCF). Because wealth measures in the SCF are the household level, they are not directly comparable to the model simulations which are at the individual level. Instead, I compare the evolution of the wealth to income ratios over the lifecycle. Total wealth (or total net worth in the SCF) is defined as the sum of all assets net of all outstanding debts but excluding the expected value of future defined benefit and social security income. For comparability with the estimation sample, I restrict the SCF sample to households in which either the head or the spouse has any type of account-based pension plan in their current job. The evolution (and heterogeneity) in the observed wealth to wage income ratios are reported in Figure 13 alongside the model predictions. Overall the model predicts well the evolution of wealth accumulation over the lifecycle, but tends to underestimate wealth accumulation at the top. This result is not surprising given that, as shown by De Nardi and Fella (2017), standard lifecycle models cannot generate the right tail of the wealth distribution without introducing nonhomothetic preferences for bequests, entrepreneurship, and medical-expense risk.

Figure 13: Ratio of net wealth to earnings by age: data vs. model

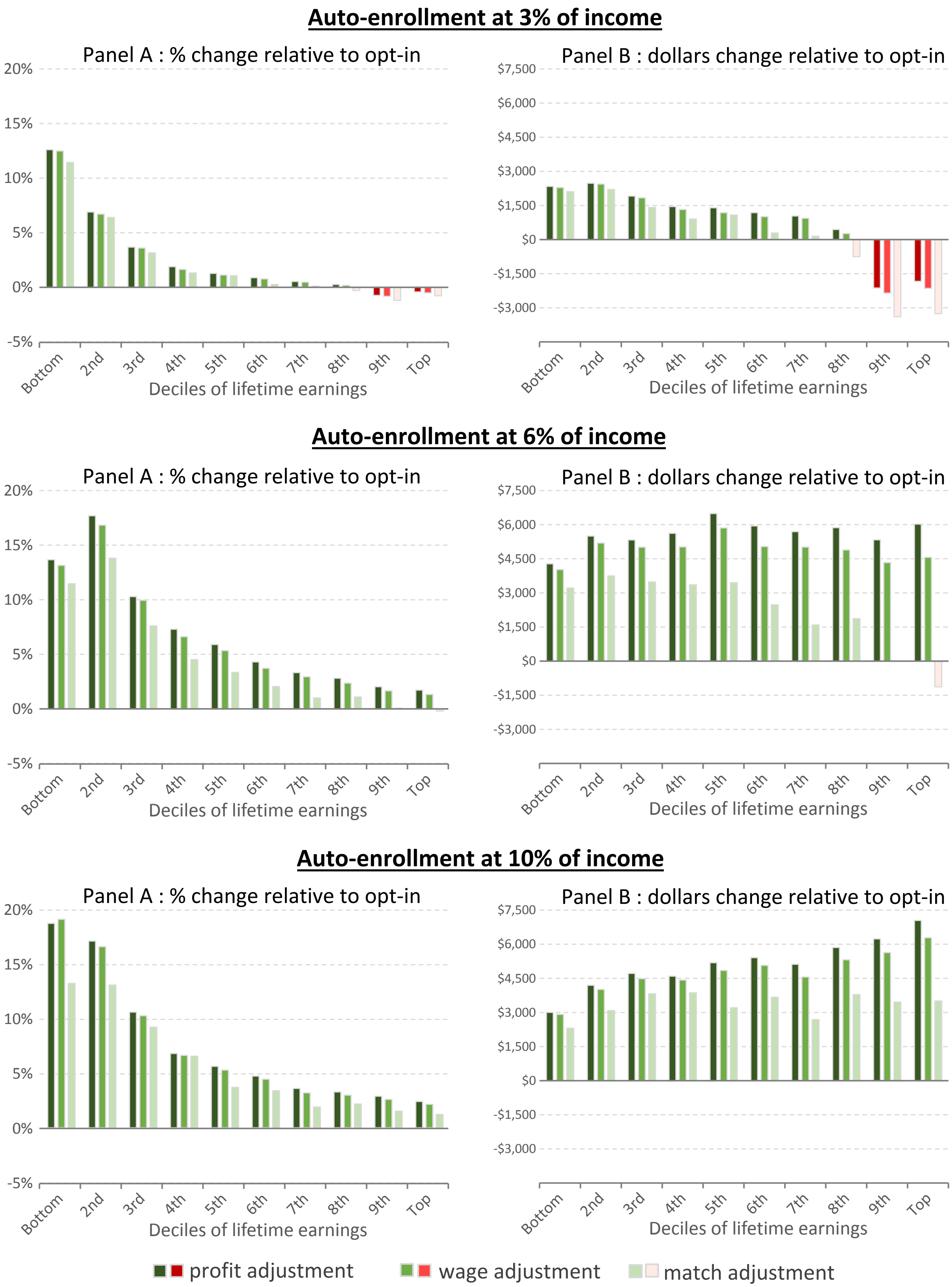


Notes: The data series (straight lines) plot the ratio of total wealth to wage income by age of the head of household. Total wealth corresponds to the sum of all assets net of all outstanding debt. The sample is restricted households where either the head or spouse has any type of account-based pension plan on a current job. The model series (dashed lines) corresponds to model simulation under an opt-in regime in all jobs. Total wealth in the model series corresponds to the sum of retirement and liquid assets net of unsecured debt liabilities. Data: Survey of Consumer Finances 2016.

C.3 Alternative Default Contribution Rates

I use the estimated model to predict the effect on lifetime savings of alternative auto-enrollment default contribution rates adopted by all employers: 3%, 6%, and 10% of income. The simulation exercise is similar to the one described in Section 6. For each auto-enrollment default I consider three different assumption about the policy incidence (see Section 8.2.2 for details). Across all counterfactuals, the effect of auto-enrollment on wealth at retirement are largest (in relative terms) for low income workers and negligible for the right tail of the income distribution. While typical auto-enrollment policies at 3% of income have a negligible effect at the median, a higher default contribution rate of 6% or 10% of income can raise median wealth at 65 by up to 5 percentage points.

Figure 14: The effect of auto-enrollment on total wealth at age 65



Notes: Each bar corresponds to the model-predicted difference between the sum of retirement and liquid wealth at age 65 under an auto-enrollment policy adopted by all employers versus an opt-in regime adopted by all employers. Each of the three series assumes a different mean of balancing employers' budget. Changes are expressed as a fraction of wealth at age 65 under the opt-in regime (panel A) or in 2006 dollars (panel B).

D U.S. Calibration - Additional Details

D.1 Labor market parameters

Data. I use the Survey of Income and Programs and Participation (SIPP) to estimate of the wage earnings process and labor market transitions probabilities. Each individual in the panel reports data on income, employment status and his or her employer identity for each week. I use the 1996 panel of the SIPP which contains data from December 1995 to February 2000 and aggregate the data at quarterly frequency. I focus on an individual's primary job (defined as the job where he worked the most hours). I restrict the sample to individuals aged 22 to 65 years old, and exclude full-time students and business owners. Following Borella et al. (2018), I use data on both male and female respondents. I use the publicly available replication files provided by Menzio et al. (2016) to build the SIPP panel and use a broadly similar definition of labor market transition variables. I assign employment status based on individuals' responses in the first week of each quarter. An individual is classified as employed if she reports having a job. I record a job-to-job transition if the identity of an individual's employer is different in two successive quarters. I record a job separation if an individual is employed in the beginning of a quarter, and not employed in the beginning of the subsequent quarter. Job separations include early retirement decisions, before the age of 65.

Earnings process. I estimate the labor earnings process for workers staying in the same job using a standard two-step minimum distance approach similar to Guvenen (2009) and Low et al. (2010). The empirical income process is given in equation (11). It is the empirical counterpart of the model earning process in equation (2) with one additional term: serially uncorrelated measurement error $\iota_{i,t} \sim N\left(0, \sigma_\iota^2\right)$.

$$\ln w_{i,t} = \delta_0 + \delta_1 a_{i,t} + \delta_2 a_{i,t}^2 + \delta_3 a_{i,t}^3 + \underbrace{\eta_{i,t}}_{\theta_{i,t}+\iota_{i,t}} \tag{11}$$

The estimation has two steps. In the first step, I estimate the parameters of the deterministic component of earnings $\left(\{\delta_j\}_{j=0}^3\right)$—a cubic in age. In the second step, I use the residual from regression (11) to estimate the five parameters governing the stochastic component of earnings: the coefficient of autocorrelation in earnings shocks (ρ), the variances of the first earnings innovation ($\sigma_{\xi_1}^2$), the variance of subsequent innovations (σ_ξ^2), and the variance of measurement error (σ_ι^2). I estimate these five parameters by minimizing the distance between the empirical variance-covariance matrix of earnings residuals and its theoretical counterpart implied by the statistical model. The parameters of the income process I estimate (given in Table 11) are in line with other results from the literature. In particular, I estimate very persistent income innovations, almost following a unit root, with a coefficient of autocorrelation $\rho = 0.974$.

Table 11: Earnings process estimates

Age component				Stochastic component of earnings			
δ_0	δ_1	δ_2	δ_3	ρ	$\sigma^2_{\xi_1}$	σ^2_{ξ}	σ^2_{ι}
1.632	0.094	-0.00107	$3.9e^{-7}$	0.974	0.184	0.0125	0.10

Notes: Quarterly earnings process estimated using a two-step minimum distance estimator on a panel of workers continuously employed in the same job. Data source: U.S. Survey of Income and Program Participation.

Earnings after a transition. I estimate the median change in log salary following a job-to-job transition (μ^{JJ}) to be equal to 0.048. I estimate that job transitions following a period of unemployment are associated with a loss in earnings. I estimate the median change in log salary relative to the last salary prior to unemployment (μ^{UE}) to be equal to -0.078.

Numeraire. The average net compensation per worker in the U.S. was around $\$37,078$ in 2006 (from the Social Security Administration national average wage index). This is also almost equal to the median annual salary in the estimation sample ($\$37,998$ in 2006 dollars).[48] I calibrate average quarterly earnings equal to $\overline{w} = \$9,250$.

Labor transition probabilities. I use SIPP micro-data to estimate quarterly job-to-job (π^{JJ}) and job to non-employment (π^{EU}) transition probabilities by age and tenure and job finding rates (π^{UE}) by age. The initial unemployment rate is set equal to 22%, which is the share not employed at age 22 in SIPP. The probability that an employed individual switches to another job (given in equation (12)) or moves to non-employment (given in equation (13)) is the sum of an age component (i.e. a fifth-order polynomial in age) and a tenure component (a set of dummies for employees in their first 9 quarters of tenure):

$$\pi^{JJ}(a, ten) = \sum_{k=1}^{5} \alpha_k^{JJ} a^k + \sum_{j=1}^{9} \iota_k^{JJ} \mathbb{1}_{(ten=j)} \tag{12}$$

$$\pi^{EU}(a, ten) = \sum_{k=1}^{5} \alpha_k^{EU} a^k + \sum_{j=1}^{9} \iota_k^{UE} \mathbb{1}_{(ten=j)} \tag{13}$$

The probability that an unemployed individual finds a job, given in equation (14), is defined as

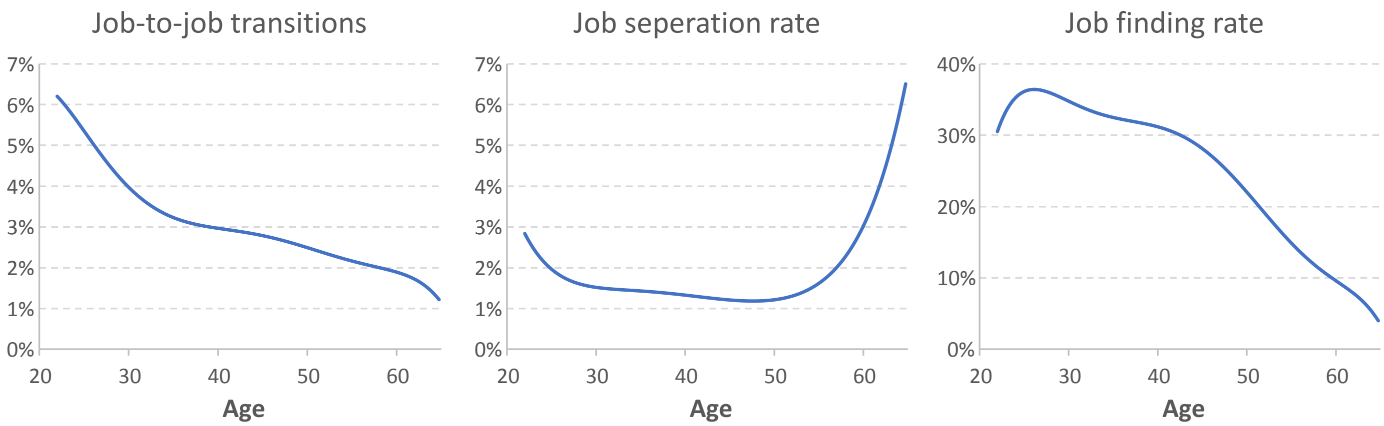
[48] I report the median rather than the average salary because the 401(k) administrative data I use do not have income data consistently for non-participants. I restrict the sample to employees hired after auto-enrollment (where participation is above 94%) and assume that the 6% of non-participants earn below the median (consistent with the fact that participation tends to be lower among low income workers).

a fifth-order polynomial in age.

$$\pi^{UE}(a) = \sum_{k=1}^{5} \alpha_k^{EU} a^k \tag{14}$$

I estimate equations (12), (13), and (14) using a linear probability regression. Estimates for the age component of labor market transitions are reported in Figure 15. Estimates for the tenure component are reported in Figure 16.

Figure 15: Age component of quarterly labor market transitions



Notes: The graphed series plot the predicted age profile of labor market transitions implied the a linear probability regression where the dependent variable is dummy for labor market transitions. The series are constructed using the coefficients of the fifth-order polynomial in age in equations (14) and (12).

Figure 16: Tenure component of quarterly labor market transitions

Notes: The graphed series show coefficients from a linear probability regression where the dependent variable is a dummy for labor market transitions. The coefficients correspond to the tenure dummies in equations (14) and (12).

D.1.1 Demographics

Survival probabilities. Survival probabilities for each age are calibrated to the U.S. Social Security 2015 Actuarial Life Tables.[49]

Equivalence scale. Changes in household composition over the lifecycle are captured by an equivalence scale in the utility function. I use the equivalence scale by age estimated by Lusardi et al. (2017). Using PSID data from 1984 to 2005, Lusardi et al. (2017) estimate $z(j_t, k_t) = (j_t + 0.7k_t)^{0.75}$ where j_t and k_t are, respectively, the average number of adults and children (under 18 years old) in a household with a head of age t. They normalize this measure by $z(2,1)$—the composition of a household with 2 adults and 1 child—to get the equivalence scale at age t equal to $n_t = \frac{z(j_t,k_t)}{z(2,1)}$. To estimate n_t I use publicly available replication files from Lusardi et al. (2017) and aggregate the data across education groups.

D.1.2 Tax and benefit system

Income taxation. Taxable income is defined as the sum of labor earnings, social security and unemployment benefits, DC withdrawals, less contributions to the DC account:

$$y_t^{tax} = \begin{cases} (1-s_t)\, w_t & if\; emp_t \in \{E, JJ\} \\ ui_t + draw_t dc_t & if\; emp_t = U \\ ss\left(ae_{A^{ret}}\right) + draw_t dc_t & if\; emp_t = Ret \end{cases}$$

Individuals' income tax liability is calculated according to the federal income tax schedule of 2006 (the first year of data and the base year for the calibration) for an individual filling as single and claiming the standard deduction. The tax formula has 5 annual income brackets $\{\tilde{\kappa}_i^{\tau}\}_{i=1}^{5} = \{\$\,5,150\,;\ \$\,7,550\,;\ \$\,30,650\,;\ \$\,74,200\,;\ \$\,154,800\}$.[50] Quarterly tax brackets are defined as: $\kappa_i^{\tau} = \frac{1}{4}\tilde{\kappa}_i^{\tau}$. The quarterly income tax liability is equal to:

[49] In teh data, death probabilities are given at a yearly frequency while individuals make decisions every quarter in the model. I assume that the quarterly death probability is equal to 0.25 times the annual probability.

[50] Note that the first bracket correspond to the standard deduction amount in 2006.

$$
tax_t^i = \begin{cases} 0 & if\, y^{tax} \leq \kappa_1^{\tau} \\ 0.10\left(y^{tax} - \kappa_1^{\tau}\right) & if\, \kappa_2^{\tau} \geq y^{tax} > \kappa_1^{\tau} \\ 0.10\left(\kappa_2^{\tau} - \kappa_1^{\tau}\right) + 0.15\left(y^{tax} - \kappa_2^{\tau}\right) & if\, \kappa_3^{\tau} \geq y^{tax} > \kappa_2^{\tau} \\ 0.10\left(\kappa_2^{\tau} - \kappa_1^{\tau}\right) + 0.15\left(\kappa_3^{\tau} - \kappa_2^{\tau}\right) + 0.25\left(y^{tax} - \kappa_3^{\tau}\right) & if\, \kappa_4^{\tau} \geq y^{tax} > \kappa_3^{\tau} \\ 0.10\left(\kappa_2^{\tau} - \kappa_1^{\tau}\right) + 0.15\left(\kappa_3^{\tau} - \kappa_2^{\tau}\right) + 0.25\left(\kappa_4^{\tau} - \kappa_3^{\tau}\right) + 0.28\left(y^{tax} - \kappa_4^{\tau}\right) & if\, \kappa_5^{\tau} \geq y^{tax} > \kappa_4^{\tau} \\ 0.10\left(\kappa_2^{\tau} - \kappa_1^{\tau}\right) + 0.15\left(\kappa_3^{\tau} - \kappa_2^{\tau}\right) + 0.25\left(\kappa_4^{\tau} - \kappa_3^{\tau}\right) + 0.28\left(\kappa_5^{\tau} - \kappa_4^{\tau}\right) + 0.33\left(y^{tax} - \kappa_5^{\tau}\right) & if\, y^{tax} > \kappa_5^{\tau} \end{cases}
$$

Public pension. The amount of public pension benefit (ss) is computed according the 2006 Social Security formula at the full retirement age , with an income floor guaranteed by the Supplemental Security Income program (with a monthly benefit $si = \$603$). Quarterly public pension benefits are equal to:

$$
ss\left(ae_{A^{ret}}\right) = 3 \times \max\left\{si\,;\, \tilde{ss}\left(ae_{A^{ret}}\right)\right\}
$$

where $\tilde{ss}$, the monthly social security benefit, is increasing in average lifetime earnings $ae_{A^{ret}}$ up to a maximum monthly benefit:

$$
\tilde{ss} = \begin{cases} 0.90 \times \frac{1}{3} ae_{A^{ret}} & if\, \frac{1}{3} ae_{A^{ret}} \leq \$656 \\ 0.90 \times \$656 + 0.32 \times \left(\frac{1}{3} ae_{A^{ret}} - \$656\right) & if\, \$3,955 > \frac{1}{3} ae_{A^{ret}} > \$655 \\ \min\left\{0.90 \times \$656 + 0.32 \times \$3,299 + \left(0.15 \times \frac{1}{3} ae_{A^{ret}} - \$3,299\right)\,;\, \$2,053\right\} & if\, \frac{1}{3} ae_{A^{ret}} > \$3,955 \end{cases}
$$

Unemployment benefits. Unemployment insurance provides a constant replacement rate ω of labor earnings implied by the labor productivity level in the last period of employment. Labor productivity θ_t stays constant during an unemployment spell. I set $\omega = 0.40$, which is the average replacement rate across all U.S. states (U.S. Department of Labor, 2018). In the U.S. unemployment benefits may be reduced by a claimant's retirement income. The Unemployment Compensation Amendments of 1976 require that all retirement income be offset against unemployment compensation. There are differences in how states implement the offset and generally, for withdrawals from a 401(k), only the amount contributed by the employer offsets unemployment benefits (see Franco (2004) for differences across U.S. states). For simplicity, I assume that the employer contribution portion of an early withdrawal is always equal to the employer match rate. This simplifying assumption is valid assuming participants contribute below the matching threshold and contributions are fully vested. Adjusted unemployment benefits for an individual unemployed since period $t - x$ are given by:

$$ui_t = \max\{0\,;\,\omega w_t(\theta_{t-x}) - draw_t dc_t \times match_e\}$$

D.1.3 Assets

Asset returns and limits. I set the annual (riskless) rate of return at 3%, which is equal to the long-term nominal yield on government bonds over the sample period between 2006 to 2017.[51] The annual interest rate on unsecured credit is set at $R^{cc} = 11.52\%$, the value estimated by Laibson et al. (2017) for the interest rate on credit card debt adjusted for both bankruptcy and inflation. The borrowing constraint prior to retirement is fixed and set equal to 74% of average quarterly earnings ($\underline{l}_{a<A^{ret}} = 0.74\overline{w}$). This is the average credit card limit estimated by Kaplan and Violante (2014) in the Survey of Consumer Finances. Unsecured borrowing is not allowed in retirement: $\underline{l}_{a\geq A^{ret}} = 0$.[52]

Asset taxation. In line with IRS rules for 2006, the maximum contribution limit for tax-deferred retirement contributions ($limit_a$) is set equal to \$3,750 per quarter (\$15,000 annually) for individuals younger than 50 years old and \$5,000 after that in 2006 dollars. The tax penalty for early DC withdrawals (pen_t) is equal to 10% before age 55 and to zero afterwards[53].Capital income is taxed at a rate τ^k of 15% which implies that the (net of tax) annual returns on liquid wealth is 2.55%.

Vesting schedule. An employee who separates from her employer before the end of the vesting period may lose part (or all) of the employer matching contribution. A vesting schedule $vst_e(ten)$ determines the percentage of employer contributions that an employee keeps if she separates at a given tenure level. For instance, under a cliff vesting schedule, vst_e is equal to zero before the end of the vesting period and one afterward. Modeling the vesting schedule explicitly would introduce an additional continuous state variable to the dynamic problem: the amount of non-vested of DC wealth. Instead, I adjust employer contributions by a factor $\Upsilon_e(t, ten)$ proportional to the risk of losing unvested employer contributions. The adjustment factor $\Upsilon_e(t, ten)$ is given in equation (15). It depends on both the cumulative job-separation probability and the vesting schedule. It is smaller than one and increasing in tenure before the end of the vesting period, and equal to one

[51] R is assumed to be exogenous, and I abstract form the equilibrium determination of the interest rate for two reasons. First, the effect of auto-enrollment policies on aggregate savings is small, and second, the effect of the policy is concentrated on low-income individuals (who hold a relatively small share of aggregate assets).

[52] This assumption is similar to Kaplan and Violante (2014). Absent this no-borrowing-in-retirement constraint, older individuals in the model borrow excessively because their high mortality risk reduces their repayment liability.

[53] In the model, early withdrawals are only allowed in periods of unemployment. The tax code allows penalty-free 401(k) hardship withdrawals for unemployed people older than 55, which is earlier than the normal 59½ eligibility age for penalty-free withdrawals.

afterward. Importantly, this specification captures the fact that vesting matters more for employees who—based on their age and tenure—are more likely to separate from their employer.

$$\Upsilon_e\left(t, ten\right) = 1 - \sum_{j=0}^{T^R-t} \left(\prod_{k=1}^{j-1} \left(1 - \pi_{t+k,ten+k}^{EU} - \pi_{t+k,ten+k}^{JJ}\right) \right) \left(\pi_{t+j,ten+j}^{EU} + \pi_{t+j,ten+j}^{JJ}\right) \left(1 - vst_e\left(ten + j\right)\right) \tag{15}$$

Employer contributions adjusted for 'vesting risk' ($\widetilde{ec}_e$) are given by equation (16).

$$\widetilde{ec}_{e,t}\left(\tau_t, ten_t\right) = \Upsilon_e\left(t, ten_t\right) \times ec_e\left(\tau_t, ten_t\right) \tag{16}$$

On average, in the estimation sample of 34 U.S. 401(k) plans, 52% of matching contribution are vested immediately and this share increases over tenure. The average vested share reaches 70% by the end of the second year of tenure. I assume that all matching contributions are fully vested starting from the 9th quarter of tenure.

E U.K. National Auto-enrollment Policy and Calibration

E.1 U.K. national auto-enrollment policy

The U.K. Pension Act of 2008 requires employers to automatically enroll their employees (with the option to opt-out) into a workplace pension scheme. The law sets the minimum employee default contribution rate at 1% of income and the minimum employer contribution at 1%. Each employer was given a staging date based on its number of employees. The staging dates for each employer size are reported in Table 12. Employers were required to enroll all employees aged between 22 and the stage pension age in a workplace pension plan by the staging date.[54] The auto-enrollment requirement applies to both new hires and the non-participating seasoned employees. I refer to Cribb and Emmerson (2016) for more details on the policy and its rollout.

[54] Employers could apply to postpone the implementation of auto-enrollment by up to 3 months after the staging date. Since I do not observe whether a firm applied to use postponement, I treat the staging date as binding.

Table 12: Auto-enrollment staging dates by employer size

Employer size	Auto-enrollment staging date	Employer size	Auto-enrollment staging date	Employer size	Auto-enrollment staging date
120,000+	October 1, 2012	2,000+	August 1, 2013	61+	August 1, 2014
50,000+	November 1, 2012	1,250+	September 1, 2013	60+	October 1, 2014
30,000+	January 1, 2013	800+	October 1, 2013	59+	November 1, 2014
20,000+	February 1, 2013	500+	November 1, 2013	58+	January 1, 2015
10,000+	March 1, 2013	350+	January 1, 2014	54+	March 1, 2015
6,000+	April 1, 2013	250+	February 1, 2014	50+	April 1, 2015
4,100+	May 1, 2013	160+	April 1, 2014	40+	August 1, 2015
4,000+	June 1, 2013	90+	May 1, 2014	30+	October 1, 2015
3,000+	July 1, 2013	62+	July 1, 2014		

Source: Cribb and Emmerson (2016).

E.2 Auto-enrollment effect after a job transition

In this Appendix I provide additional evidence on the effect of auto-enrollment after a job described in Section 2.4.

In the model, I interpret the difference in participation and contribution rates between those who were previously autoenrolled and those who weren't as reflecting a wealth effect (more saving in the previous job reduces the need to save in the current job). However, previously auto-enrolled wokers may be better informed about the future roll-out of the policy and expect to be eventually automatically enrolled by their new employer. They may therefore choose not to opt into their new employer's retirement savings plan because they expect to soon be automatically enrolled anyways. This mechanism should mainly affect those whose new employer will soon become subject to the auto-enrollment mandate. To test this hypothesis in Table 13 columns (3) and (6), I interact the transition to an opt-in employer with a dummy equal to 1 if the new employer is expected to become subject to automatic enrollment within the next 12 months. While the results are noisy and not statistically significant, previously autoenrolled workers who do not expect to be autoenrolled in their new job within a year reduce their participation by 7.5 percentage points and their contributions by 0.12% of income which is in line with the model predictions in 7.2 (a model-implied drop in participation of 9.6% and in contributions of 0.11% of salary). The drop in participation and contributions for those whose employer was scheduled to adopt auto-enrollment within 12 months is significantly larger (participation drops down by an additional 11.6 percentage points and contributions down by an additional 0.93% of income) which suggests that differential expectations about the policy rollout may play a role.

Table 13: Auto-enrollment effect after a job transition

	Participation rate			Contribution rate		
	(1)	(2)	(3)	(4)	(5)	(6)
AE to non-AE employer	-0.065**	-0.128**	-0.075	-0.267**	-0.538**	-0.122
	(0.031)	(0.055)	(0.052)	(0.108)	(0.227)	(0.266)
AE to non-AE employer ×			-0.116			-0.927*
AE to be adopted within a year			(0.092)			(0.517)
$\text{Size}_{e'} \times \text{Size}_e$		✓	✓		✓	✓
Employee characteristics	✓	✓	✓	✓	✓	✓
Employer×Year FE	✓	✓	✓	✓	✓	✓
Observations	37120	37120	37120	37120	37120	37120

* p<0.10, ** p<0.05, *** p<0.01

Notes: Each cell reports coefficients from a regression whose dependent variable is a dummy for workers' participation status in the retirement savings plan (columns 1, 2 and 3), or contribution rate in percentage points of pensionable pay (columns 4, 5, and 6). The sample contains job switchers in their first year of tenure. I assign an enrollment regime to each employer (i.e. AE or non-AE) based on the U.K. national auto-enrollment policy rollout schedule (in Appendix E). The coefficients for "AE to non-AE employer" correspond to the effect of the automatic enrollment mandate in the previous employer on contributions when the current employer is not subject to the auto-enrollment requirement. In Columns 3 and 6, the variable "AE to be adopted within a year" is a dummy equal to 1 if the current employer is expected to be subject to automatic enrollment in the 12 months following the date of observation. Controls for employees' characteristics include controls for log salary and log tenure in the current job, log salary and log tenure when last observed in the previous job, a dummy for gender, and a third order polynomial in age. The sample is restricted to private sector workers between the ages of 22 and 60. Standard errors clustered by current employer are in parentheses. Data source: U.K. Annual Survey of Hours and Earnings waves 2006 to 2017.

E.3 U.K. calibration

In what follows I detail the parametrization of the model in the U.K. validation exercise presented in Section 7.2. Other parameters are set equal to their value in the U.S. calibration.

Employer matching formulas. ASHE does not collect data on each employer's matching formula.[55] I use data on employee and employer contributions prior to the auto-enrollment rollout to back out each employer's matching formula. The distribution of contribution rates often features bunching at the threshold on employer matching. I hypothesize that the modal (positive) contribution rate in a given firm is the threshold on employer matching. I test this hypothesis in Table 14 by regressing employer contributions on employee contributions above and below the modal contributions. Columns 1 and 2 show that, up to the modal employee contribution, each percentage point increase in contributions raises the employer contribution by half a percentage point. Columns 1 and

[55] I am not aware of any representative evidence on the distribution of employer matching formulas in the U.K.

3 show that employee contributions above the modal contribution rate do not lead to significantly higher employer contributions. These results are consistent with employers offering a 50% match on contributions up to the modal contribution.

In the calibration of the model to the U.K. environment, I group employers into 5 types based on their employer contribution formula. Motivated by the evidence in Table 14, I set the match rate equal to 50% for all employers with immediate vesting. I assume that employers have one of five different matching thresholds (cap_e), based on the five most common modal contribution rates observed in the data (and covering more than 80% of employees). The calibrated probability distribution of employer types is reported in Table 15 and reflects the empirical distribution of modal employee contributions.

U.K. income process. I estimate the income process using earnings data for U.K. private sector workers aged 22 to 65 using ASHE data between 1997 and 2016. I use the same two step procedure used to estimate the income process for U.S. workers. Estimation results are reported in Table 16. I adjust the estimate for the fact that AHSE is collected at annual frequency while the model is simulated at quarterly frequency.[56]

[56] Given the auto-correlation ($\hat{\rho}$) and variance ($\hat{\sigma}_\xi^2$) estimated using annual data, the quarterly values are given by $\rho = \hat{\rho}^{\frac{1}{4}}$ and $\sigma_\xi^2 = \frac{1}{\left(1+\rho^2+\rho^4+\rho^6\right)}\hat{\sigma}_\xi^2$

Table 14: Relationship between employer and employee retirement contributions in the U.K.

	Employer contribution in % of pay		
	(1)	(2)	(3)
	Full sample	Employee contrib. $\leq$ modal contrib.	Employee contrib. $>$ modal contrib.
Employee contribution rate			
$\times$dummy (contrib.$\leq$modal)	0.500***	0.504***	
	(0.14)	(0.14)	
$\times$dummy (contrib.$>$modal)	0.229		0.155
	(0.158)		(0.123)
Modal contribution rate in the firm			
$\times$dummy (contrib.$\leq$modal)	0.471	0.499	
	(0.306)	(0.333)	
$\times$dummy (contrib.$>$modal)	0.923**		0.812**
	(0.408)		(0.407)
R-squared	0.146	0.143	0.111
Observations	21771	17856	3915

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: The outcome variable is the employer retirement contribution amount as a percentage of pensionable pay. The modal (positive) employee contribution is computed separately for each employer. I restrict the sample to firms with at least 50 observations in the three years prior to the auto-enrollment staging date. Data source: ASHE 2006 to 2016.

Table 15: Probability distribution of employer types

	Threshold on matching (cap_e)				
	2%	3%	4%	5%	6%
Prob.	0.12	0.18	0.15	0.25	0.30

Notes: This table shows the probability distribution of match thresholds in the U.K. calibration of the model. This distribution is inferred from the empirical distribution of modal contribution rates in the three years prior to the auto-enrollment staging date. Data source: ASHE 2006 to 2016.

Table 16: U.K. earnings process estimates

Age component				Stochastic component of earnings			
δ_0	δ_1	δ_2	δ_3	$\hat{\rho}$	$\sigma^2_{\xi_1}$	$\hat{\sigma}^2_{\xi}$	σ^2_m
5.306	0.255	-0.0043	0.0000219	0.949	0.2447	0.0232	0.2442

Notes: Annual earnings process estimated using a two-step minimum distance estimator on a panel of workers continuously employed in the same job. Data source: ASHE 1997 to 2016.

Early withdrawals. Early withdrawals from a DC account are not allowed in the U.K. Thus, I impose a 100% tax penalty on early withdrawals $pen_a = 1$.

Income taxation. Individuals' income tax liability is calculated according to the U.K. income tax schedule of 2006 for an individual claiming the personal allowance for singles. The quarterly income tax liability is equal to:

$$tax_t^i = \begin{cases} 0 & if\ 4y^{tax} \leq £4,895 \\ 0.22\left(y^{tax} - \frac{4,895}{4}\right) & if\ £32,400 \geq 4y^{tax} > £4,895 \\ 0.22\left(\frac{32,400-4,895}{4}\right) + 0.40\left(y^{tax} - \frac{32,400}{4}\right) & if\ 4y^{tax} > £32,400 \end{cases}$$

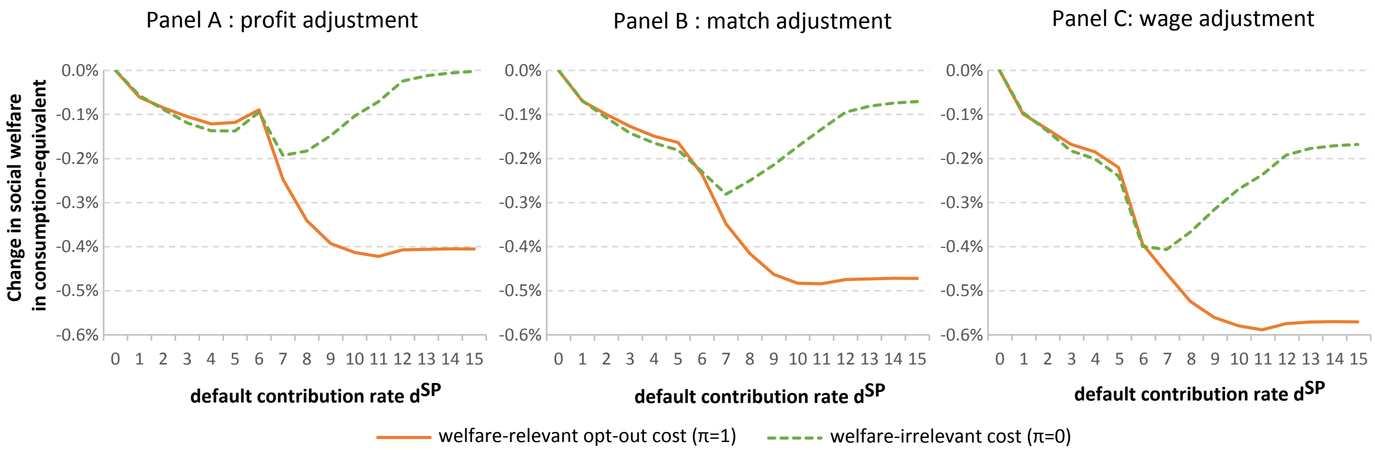
Public pension. The U.K. State Pension provides public pension benefits similar to Social Security in the U.S. The relationship between average lifetime earnings and public pension entitlements is modeled to match the evidence in Figure 1 of O'Dea (2018), expressed in 2006 pounds.

Unemployment benefits. The main unemployment benefit in the U.K. is the Jobseeker's Allowance. In 2006, the quarterly allowance was equal to £ 746 for unemployed individuals older than 24.

F Welfare Results

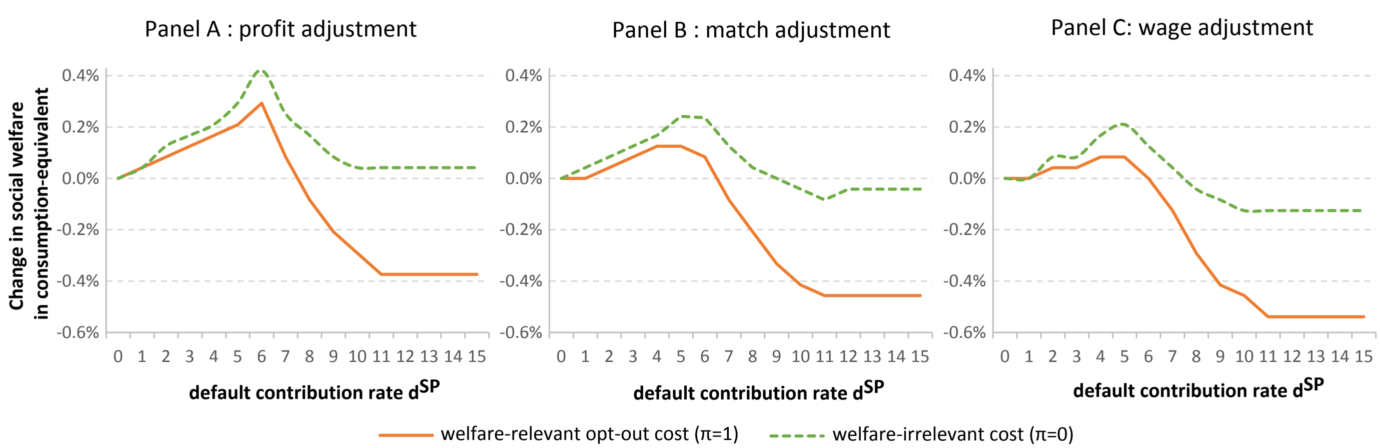
In this Appendix, I present the change in lifetime welfare for different default contribution rates under alternative assumptions on social preferences and on the policy incidence (as discussed in Section 8). The optimal policies reported in Table 5 represent the highest-value points in Figures 17, 18, and 19.

Figure 17: Welfare effect of alternative defaults - Utilitarian policymaker



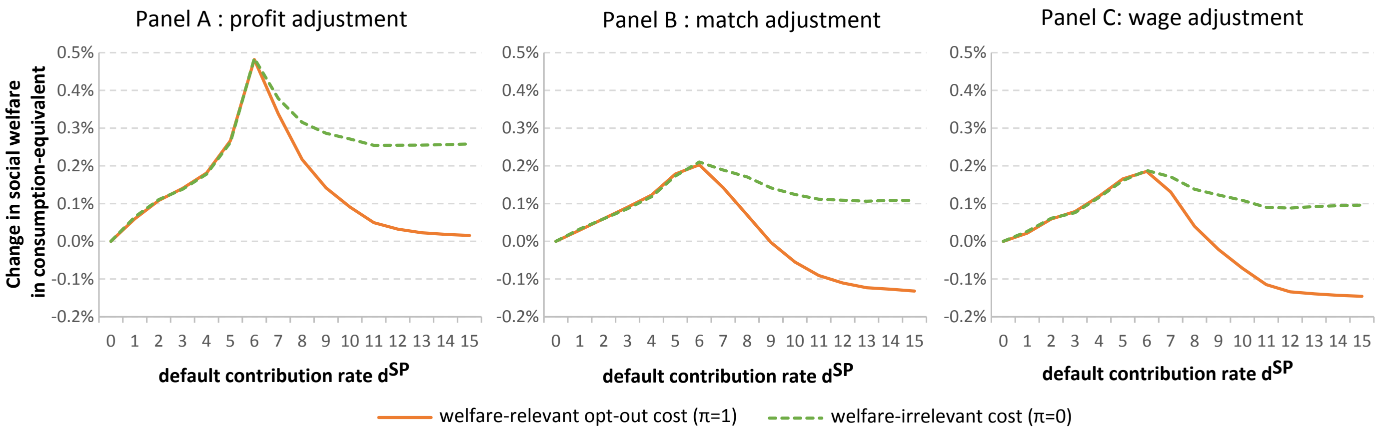
Notes: Each series correspond to social welfare under alternative default contribution rates adopted by all employers over a lifetime. Welfare is expressed in consumption-equivalent variation relative to an opt-in regime (i.e. $d^{SP} = 0$). Social preferences are assumed to be equal to individuals' preferences ($\delta^{SP} = \hat{\delta}$ and $\varphi_t = 1$).

Figure 18: Welfare effect of alternative defaults - Inequality-averse policymaker



Notes: Each series correspond to social welfare under alternative default contribution rates adopted by all employers over a lifetime. Welfare is expressed in consumption-equivalent variation relative to an opt-in regime (i.e. $d^{SP} = 0$). Social and individual time preferences are identical ($\delta^{SP} = \hat{\delta}$). Low-income individuals receive higher Pareto weights φ_t, and these weights are kept constant across policy counterfactuals.

Figure 19: Welfare effect of alternative defaults - Paternalistic policymaker



Notes: Each series correspond to social welfare under alternative default contribution rates adopted by all employers over a lifetime. Welfare is expressed in consumption-equivalent variation relative to an opt-in regime (i.e. $d^{SP} = 0$). Social preferences do not discount the future ($\delta^{SP} = 1$), and there is no redistributive motive beyond declining marginal utility ($\varphi_t = 1$).

G Numerical procedure

Discretization. I have five continuous state variables that need to be discretized: labor productivity, tenure, liquid assets, retirement wealth, and average lifetime earnings. Labor productivity is placed on a grid with 7 elements using the method of Tauchen (1986). Tenure is placed on a grid with 9 elements (with all periods past the 9th quarter of tenure treated equally). Liquid assets and retirement wealth are discretized in a way that gives smaller gaps between successive entries on the grid at lower levels. For each age, liquid assets are placed on a grid with 20 elements (7 for strictly negative values and 12 for strictly positive values), and retirement wealth on a grid with 16 elements. Average earnings are placed on a grid of 7 elements.

Consumption is not placed on a grid, and individuals can choose any feasible consumption level. 401(k) contribution rates are naturally discrete, and in the firms in the estimation sample contributions are often restricted to be below 15% of salary. Thus, I adopt a grid of contribution rates with 16 elements: 0%, 1%,...,15% of salary. Early withdrawals from the DC account are restricted to take one of ten values. That is, unemployed individuals can withdraw 0%, 1%, 2%, 5%, 10%, 15%, 25%, 50%, 75% or 100% of their DC wealth. In retirement, DC withdrawals can take one of 101 values (0%, 1%,2%,...,99%, 100%).

Interpolation. In order to evaluate the value function at points not in the discrete subset of points in the discretized state space, I use linear interpolation in multiple dimension. To limit approximation error despite having fairly coarse grids, I use a method proposed by Carroll (2012). I approximate a quasi-linear transformation of the value function. The transformed value function is closer to linear than the actual one, which implies that linear interpolation is more accurate despite having few grid points. I verify the accuracy of my numerical procedure by simulating the model—at the parameter values in my baseline estimation—with more grid points and obtain similar results.

Optimization. I solve for the DC contribution and withdrawal choices using a grid search. Given a DC contribution or withdrawal choice, I solve for the liquid asset level (which implies the consumption level) using a golden section search.

Estimation. I estimate the model in three steps. Each step involves solving the model for 2,744 unique combinations of the model's three parameters. After each step, I reduce the step-size and center the grid of parameters under consideration around the optimal value found in the previous step. For each combination of preference parameters, I solve for the value function by iterating the problem from the last period of life (period 272 corresponding to age 90). I then simulate 5,000 individuals in the opt-in group and 5,000 in the auto-enrollment group with identical realization of the stochastic variables. Estimation is computationally intensive. The program was compiled with Intel Fortran and each estimation round was parallelized on 392 computer processors using Yale High Power Computing facility.

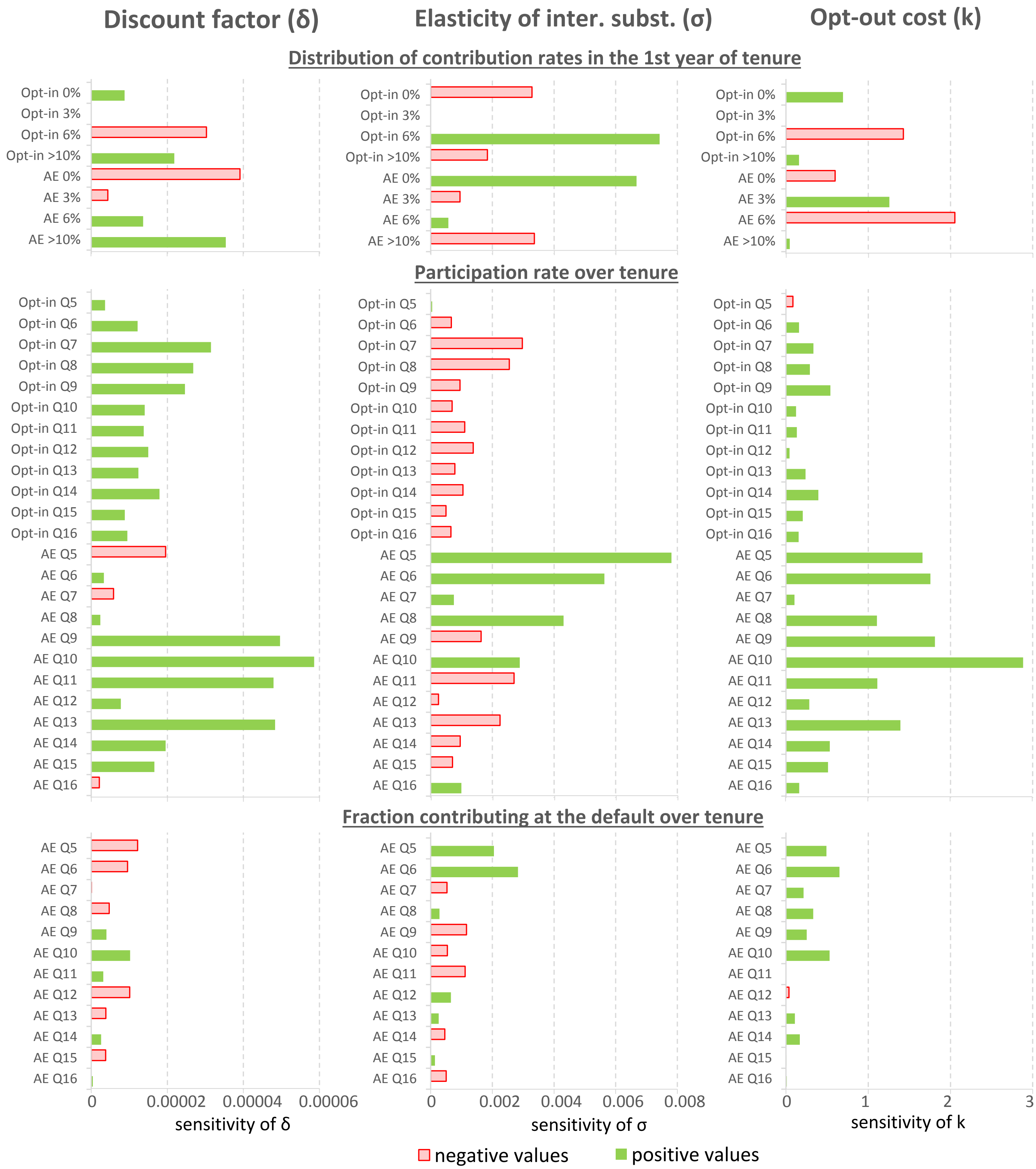
H Sensitivity

I plot in Figure 20 the elements of the sensitivity matrix as defined in Andrews et al. (2017). These values capture how a 1 percentage point increase in each empirical moment changes the estimated parameters.

Time preferences. According to the sensitivity matrix plotted in Figure 20, more contributions at the employer match threshold (6% of salary), and fewer contributions above (≥10%), would imply to a higher estimate of the elasticity of intertemporal substitution. In contrast, for the auto-enrolled, both contributions at the match threshold and above increase the estimated discount factor. This result is consistent with the finding in Best et al. (2018) that bunching at—with missing mass above—an interest rate notch can identify the elasticity of intertemporal substitution separately from the discount factor.

Opt-out cost. The sensitivity measure implies that higher participation under auto-enrollment would increase the estimated opt-out cost. Note, however, that there is limited potential for a higher opt-out cost estimate because participation under auto-enrollment is above 95% in the data.

Figure 20: Sensitivity of preference parameters to estimation moments



Notes: This figure shows the sensitivity matrix as defined in Andrews et al. (2017). Sensitivity values are rescaled to correspond to a 1 percentage point increase in each moment.

I Proofs of Section 2.2

Let $\mathrm{V_t^s}(\mathrm{d})$ be the value of contribution s when the default is equal to d:

$$\mathrm{V_t^s}(\mathrm{d}) = U(s, d) + \delta V_{t+1}(s)$$

I consider two default contribution rates $\underline{d}$ and $\overline{d}$ such that $\underline{d} < \overline{d}$. I divide the contribution choice set S into contribution rates strictly below $\underline{d}$ ($\underline{S} = \{s < \underline{d}\}$) and contribution rates above ($\overline{S} = \{s \geq \underline{d}\}$). A worker facing a default d chooses a contribution rate s^* such that:

$$\mathrm{s}^* = \underset{\mathrm{s \in S}}{\mathrm{argmax}} \left\{ \max \{\mathrm{V_t^s}(\mathrm{d})\}_{\mathrm{s \in \underline{S}}}, \max \{\mathrm{V_t^s}(\mathrm{d})\}_{\mathrm{s \in \overline{S}}} \right\}$$

In what follows, I show that a model with opt-out costs and models with either loss-averse preferences or psychological anchoring contributions, make opposite predictions about the effect of increasing the default from $\underline{d}$ to $\overline{d}$ on the likelihood of contribution strictly below $\underline{d}$.

I.1 Opt-out costs

Proof. In what follows I provide a demonstration of Proposition 1. The value function with opt-out costs is given by:

$$V_t^s(d) = \begin{cases} u((1-s) w_t - k) + \delta V_{t+1}(s) & \text{if } s \neq d \\ u((1-s) w_t) + \delta V_{t+1}(s) & \text{if } s = d \end{cases}$$

I assume that $u(.)$ is strictly increasing and concave ($u^{'} > 0$, $u^{''} < 0$), and that the continuation value $V_{t+1}(.)$ is strictly increasing and concave in the retirement contribution rate ($V^{'} > 0$, $V^{''} < 0$).[57]

A worker selects a contribution rate $s^* \neq d$ if it is better than staying at the default:

$$V_t^{s^*}(d) - V_t^d(d) > 0$$

[57]The assumption that V_{t+1} is concave in the retirement contribution rate may not always hold because of the discrete nature of the contribution choice decision. In practice, with sufficient uncertainty about the future, the continuation value V_{t+1} is generally concave.

and (conditional on opting out) s is preferred to all other contribution rates:

$$s^* = \arg\max \{V_t^s(s)\}_{s \in S}$$

Let $\underline{s} \in \underline{S}$ such that $\overline{d} > \underline{d} > \underline{s}$. Since u is increasing and strictly concave in take home pay:

$$\frac{u((1-\underline{s})w_t) - u((1-\underline{d})w_t)}{(\underline{d}-\underline{s})w_t} < \frac{u((1-\underline{s})w_t) - u\left((1-\overline{d})w_t\right)}{\left(\overline{d}-\underline{s}\right)w_t} \tag{17}$$

Since V is strictly increasing and concave in the retirement contribution rate:

$$\frac{V_{t+1}(\underline{d}) - V_{t+1}(\underline{s})}{(\underline{d}-\underline{s})} > \frac{V_{t+1}\left(\overline{d}\right) - V_{t+1}(\underline{s})}{\left(\overline{d}-\underline{s}\right)}$$

$$\Rightarrow -\delta\frac{V_{t+1}(\underline{d}) - V_{t+1}(\underline{s})}{(\underline{d}-\underline{s})w_t} < -\delta\frac{V_{t+1}\left(\overline{d}\right) - V_{t+1}(\underline{s})}{\left(\overline{d}-\underline{s}\right)w_t} \tag{18}$$

Combining (17) and (18):

$$\frac{(u((1-\underline{s})w_t) + \delta V_{t+1}(\underline{s})) - (u((1-\underline{d})w_t) + \delta V_{t+1}(\underline{d}))}{(\underline{d}-\underline{s})w_t} < \frac{(u((1-\underline{s})w_t) + \delta V_{t+1}(\underline{s})) - \left(u\left((1-\overline{d})w_t\right) + \delta V_{t+1}\left(\overline{d}\right)\right)}{\left(\overline{d}-\underline{s}\right)w_t}$$

$$\Rightarrow \frac{V_t^{\underline{s}}(\underline{s}) - V_t^{\underline{d}}(\underline{d})}{\underline{d}-\underline{s}} < \frac{V_t^{\underline{s}}(\underline{s}) - V_t^{\overline{d}}\left(\overline{d}\right)}{\overline{d}-\underline{s}} \tag{19}$$

Consider a worker whose preferred contribution rate, conditional on switching, $\underline{s}^*$ belongs to $\underline{S}$:

$$\underline{s}^* = \arg\max \left\{V_t^{\underline{s}}(\underline{s})\right\}_{\underline{s} \in \underline{S}} = \arg\max \{V_t^s(s)\}_{s \in S}$$

$$\Rightarrow V_t^{\underline{s}^*}(\underline{s}^*) - V_t^{\underline{d}}(\underline{d}) > 0 \quad \text{and} \quad V_t^{\underline{s}^*}(\underline{s}^*) - V_t^{\overline{d}}\left(\overline{d}\right) > 0$$

Combining this result with inequality (19) implies that for this worker:

$$V_t^{\underline{s}^*}(\underline{s}^*) - V_t^{\underline{d}}(\underline{d}) < V_t^{\underline{s}^*}(\underline{s}^*) - V_t^{\overline{d}}\left(\overline{d}\right)$$

$$\Rightarrow V_t^{\underline{d}}(\underline{d}) > V_t^{\overline{d}}\left(\overline{d}\right)$$

This implies that:

$$\Pr\left(V_t^{\underline{s}^*}\left(\underline{d}\right) > V_t^{\underline{d}}\left(\underline{d}\right)\right) \leq \Pr\left(V_t^{\underline{s}^*}\left(\overline{d}\right) > \widetilde{V}_t^{\overline{s}_1^*}\left(\overline{d}\right)\right)$$

$$\Rightarrow \Pr\left(\mathrm{s}^* \in \underline{\mathrm{S}} \,|\, \underline{\mathrm{d}}\right) \leq \Pr\left(\mathrm{s}^* \in \underline{\mathrm{S}} \,|\, \overline{\mathrm{d}}\right)$$

□

I.2 Loss aversion

Proof. In what follows I provide a demonstration of proposition (2). The value function of a loss averse agent contributing s is given by the following expression (with $v^{'} < 0$, $\eta > 0$ and $\lambda \geq 1$).

$$V_t^s\left(d\right) = U\left(s, d\right) + \beta\left(1 - m_a\right)\mathbb{E}_t\left(V_{t+1}\left(s\right)\right)$$

$$U\left(s, d\right) = \begin{cases} u_a\left(c_t\left(s\right)\right) + \eta\left(v\left(s\right) - v\left(d\right)\right) & \text{if } s < d \\ u_a\left(c_t\left(s\right)\right) + \eta\lambda\left(v\left(s\right) - v\left(d\right)\right) & \text{if } s \geq d \end{cases}$$

I define $\widetilde{V_t^s}\left(d, \vec{X}_t\right)$, a monotone transformation of $V_t^s\left(d, \vec{X}_t\right)$, such that:

$$\widetilde{V_t^s}\left(d\right) = V_t^s\left(d\right) + \eta v\left(d\right)$$

$$\Leftrightarrow \widetilde{V}_t^s\left(d\right) = u_a\left(c_t\right) + \eta v\left(s\right) + \mathbb{1}_{(s \geq d)}\eta\left(\lambda - 1\right)\left(v\left(s\right) - v\left(d\right)\right) \tag{20}$$

Since $\eta v\left(d\right)$ does not depend on s, the optimal contribution choice $s^*$is such that :

$$\mathrm{s}^* = \underset{\mathrm{s} \in \mathrm{S}}{\mathrm{argmax}}\left\{\mathrm{V}_\mathrm{t}^\mathrm{s}\left(\mathrm{d}\right)\right\} = \underset{\mathrm{s} \in \mathrm{S}}{\mathrm{argmax}}\left\{\widetilde{\mathrm{V}_\mathrm{t}^\mathrm{s}}\left(\mathrm{d}\right)\right\}$$

First, I consider the case where $s \in \underline{S}$. From (20) I get that:

$$\widetilde{V}_t^s\left(\underline{d}\right) = \widetilde{V}_t^s\left(\overline{d}\right) = u_a\left(c_t\right) + \eta v\left(s\right)$$

I define $\underline{s}^*$ such that:

$$\underline{s}^* = \underset{s \in \underline{S}}{argmax}\left\{V_t^s\left(\underline{d}\right)\right\} = \underset{s \in \underline{S}}{argmax}\left\{V_t^s\left(\overline{d}\right)\right\}$$

This implies that:

$$\Rightarrow V_t^{\underline{s}^*}\left(\underline{d}\right)=V_t^{\underline{s}^*}\left(\overline{d}\right) \tag{21}$$

I next consider the case where $s\in\overline{S}$. From (20) I get that:

$$\widetilde{V}_t^s\left(\overline{d}\right)-\widetilde{V}_t^s\left(\underline{d}\right)=\begin{cases}\eta\left(\lambda-1\right)\left(v\left(\underline{d}\right)-v\left(s\right)\right)\geq 0 & \text{if } \underline{d}\leq s<\overline{d}\\ \eta\left(\lambda-1\right)\left(v\left(\underline{d}\right)-v\left(\overline{d}\right)\right)\geq 0 & \text{if } s\geq\overline{d}\end{cases}$$

$$\Rightarrow \widetilde{V}_t^s\left(\underline{d}\right)\leq\widetilde{V}_t^s\left(\overline{d}\right)\quad \forall s\in\overline{S} \tag{22}$$

I defined $\overline{s}_1^*$ and $\overline{s}_2^*$ such that:

$$\overline{s}_1^*=\underset{s\in\overline{S}}{argmax}\left\{V_t^s\left(\underline{d}\right)\right\}\quad and\quad \overline{s}_2^*=\underset{s\in\overline{S}}{argmax}\left\{V_t^s\left(\overline{d}\right)\right\}$$

Inequality (22) implies that:

$$\Rightarrow \widetilde{V}_t^{\overline{s}_1^*}\left(\underline{d}\right)\leq\widetilde{V}_t^{\overline{s}_2^*}\left(\overline{d}\right) \tag{23}$$

Combining (21) and (23), I get that :

$$\Pr\left(V_t^{\underline{s}^*}\left(\underline{d}\right)>\widetilde{V}_t^{\overline{s}_1^*}\left(\underline{d}\right)\right)\leq\Pr\left(V_t^{\underline{s}^*}\left(\overline{d}\right)>\widetilde{V}_t^{\overline{s}_1^*}\left(\overline{d}\right)\right)$$

$$\Rightarrow \Pr\left(\mathrm{s}^*\in\underline{\mathrm{S}}\,|\,\underline{\mathrm{d}}\right)\leq\Pr\left(\mathrm{s}^*\in\underline{\mathrm{S}}\,|\,\overline{\mathrm{d}}\right)$$

□

I.3 Psychological anchoring

Proof. In what follows I provide a demonstration of proposition (3). The value function under psychological anchoring is given by the following expression (with $\chi>0$).

$$V_t^S\left(d\right)=\begin{cases}u_a\left(c_t\left(s\right)\right)+\left(\delta+\chi\right)\left(1-m_a\right)\mathbb{E}_t\left(V_{t+1}\left(d\right)\right) & \text{if } s<d\\ u_a\left(c_t\left(s\right)\right)+\delta\left(1-m_a\right)\mathbb{E}_t\left(V_{t+1}\left(d\right)\right) & \text{if } s=d\\ u_a\left(c_t\left(s\right)\right)+\left(\delta-\chi\right)\left(1-m_a\right)\mathbb{E}_t\left(V_{t+1}\left(d\right)\right) & \text{if } s>d\end{cases} \tag{24}$$

First, I consider the case where $s\in\underline{S}$. From (24) I get that:

$$V_t^s\left(\underline{d}\right) = V_t^s\left(\overline{d}\right)$$

$$\underline{s}^* = \underset{s\in\underline{S}}{argmax}\left\{V_t^s\left(\underline{d}\right)\right\} = \underset{s\in\underline{S}}{argmax}\left\{V_t^s\left(\overline{d}\right)\right\}$$

$$\Rightarrow V_t^{\underline{s}^*}\left(\underline{d}\right) = V_t^{\underline{s}^*}\left(\overline{d}\right) \tag{25}$$

Next, I move to the case where $s \in \overline{S}$. I consider 3 possibilities:

If $s = \left\{\underline{d}, \overline{d}\right\}$ then:

$$V_t^s\left(\overline{d}\right) - V_t^s\left(\underline{d}\right) = \chi\left(1 - m_a\right)\mathbb{E}_t\left(V_{t+1}\left(s\right)\right) > 0$$

If $\overline{d} > s > \underline{d}$ then:

$$V_t^s\left(\overline{d}\right) - V_t^s\left(\underline{d}\right) = 2\chi\left(1 - m_a\right)\mathbb{E}_t\left(V_{t+1}\left(s\right)\right) > 0$$

If $s > \overline{d}$ then:

$$V_t^s\left(\underline{d}\right) = V_t^s\left(\overline{d}\right)$$

I define $\overline{s}_1^*$ and $\overline{s}_2^*$ as the preferred contribution in the set $\overline{S}$ under each default:

$$\overline{s}_1^* = \underset{s\in\overline{S}}{argmax}\left\{V_t^s\left(\underline{d}\right)\right\} \quad and \quad \overline{s}_2^* = \underset{s\in\overline{S}}{argmax}\left\{V_t^s\left(\overline{d}\right)\right\}$$

Given that:

$$V_t^s\left(\underline{d}\right) \leq V_t^s\left(\overline{d}\right) \quad \forall s \in \overline{S}$$

It follows that:

$$V_t^{\overline{s}_1^*}\left(\underline{d}\right) \leq V_t^{\overline{s}_2^*}\left(\overline{d}\right) \tag{26}$$

Combining (25) and (26), I get that :

$$\Pr\left(V_t^{\underline{s}^*}\left(\underline{d}\right) > \widetilde{V}_t^{\overline{s}_1^*}\left(\underline{d}\right)\right) \leq \Pr\left(V_t^{\underline{s}^*}\left(\overline{d}\right) > \widetilde{V}_t^{\overline{s}_1^*}\left(\overline{d}\right)\right)$$

$$\Rightarrow \Pr\left(\mathrm{s}^* \in \underline{\mathrm{S}}\,|\,\underline{\mathrm{d}}\right) \leq \Pr\left(\mathrm{s}^* \in \underline{\mathrm{S}}\,|\,\overline{\mathrm{d}}\right)$$

$\square$

References

Altmann, S., F. Armin, P. Heidhues, and R. Jayaraman (2018). Defaults and donations: Evidence from a field experiment. *Review of Economics and Statistics*.

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics 132*(4), 1553–1592.

Arnoud, A., T. Choukhmane, J. Colmenares, C. O'Dea, and A. Parvathaneni (2021). The evolution of u.s. firms retirement plan offerings: Evidence from a new panel data set.

Bernheim, B. D., A. Fradkin, and I. Popov (2015). The welfare economics of default options in 401 (k) plans. *American Economic Review 105*(9), 2798–2837.

Bernheim, B. D. and A. Rangel (2009). Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *The Quarterly Journal of Economics 124*(1), 51–104.

Bernheim, D. and J. M. Gastell (2020). Optimal default options: The case for opt-out minimization.

Beshears, J., J. Choi, D. Laibson, and B. Madrian (2008). A primer on 401(k) loans.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2009). The importance of default options for retirement saving outcomes: Evidence from the united states. In *Social security policy in a changing environment*, pp. 167–195. University of Chicago Press.

Beshears, J., J. J. Choi, D. Laibson, B. C. Madrian, and W. L. Skimmyhorn (2021). Borrowing to save? the impact of automatic enrollment on debt. *Journal of Finance*.

Best, M., J. Cloyne, and E. Ilzetzki (2018). Estimating the elasticity of intertemporal substitution using mortgage notches. *Review of Economic Studies*.

BLS (2019). Number of jobs, labor market experience, and earnings growth:results from a national longitudinal survey. Technical report, Bureau of Labor Statistics.

Blumenstock, J., M. Callen, and T. Ghani (2018). Why do defaults affect behavior? experimental evidence from afghanistan. *American Economic Review 108*(10).

Borella, M., M. De Nardi, and F. Yang (2018). The aggregate implications of gender and marriage. *The Journal of the Economics of Ageing 11*, 6–26.

Brown, J. R., A. M. Farrell, and S. J. Weisbenner (2016). Decision-making approaches and the propensity to default: Evidence and implications. *Journal of Financial Economics 121*(3), 477–495.

Brown, J. R. and A. Previtero (2016). Saving for retirement, annuities and procrastination.

Bubb, R. and P. L. Warren (2018). An equilibrium theory of retirement plan design.

Butrica, B. and N. Karamcheva (2012). Automatic enrollment, employee compensation, and retirement security.

Caplin, A. and J. Leahy (2004). The social discount rate. *Journal of political Economy 112*(6), 1257–1268.

Caplin, A. and D. Martin (2017). Defaults and attention: the drop out effect. *Revue économique* (5), 747–755.

Carroll, C. D. (2012). Solution methods for microeconomic dynamic stochastic optimization problems.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. *The quarterly journal of economics 124*(4), 1639–1674.

Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from denmark. *The Quarterly Journal of Economics 129*(3), 1141–1219.

Choi, J., D. Laibson, B. C. Madrian, and A. Metrick (2006). Saving for retirement on the path of least resistance. In *Behavioral Public Finance: Toward a New Agenda.*

Choi, J. J., D. Laibson, and B. C. Madrian (2009). Reducing the complexity costs of 401 (k) participation through quick enrollment. In *Developments in the Economics of Aging*, pp. 57–82. University of Chicago Press.

Choi, J. J., D. Laibson, and B. C. Madrian (2011). $100 bills on the sidewalk: Suboptimal investment in 401 (k) plans. *Review of Economics and Statistics 93*(3), 748–763.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004). For better or for worse: Default effects and 401 (k) savings behavior. In *Perspectives on the Economics of Aging*, pp. 81–126. University of Chicago Press.

Clark, J., S. Utkus, and J. Young (2015). Automatic enrollment: The power of the default. *Valley Forge, PA: Vanguard*.

Cribb, J. and C. Emmerson (2016). What happens when employers are obliged to nudge? automatic enrolment and pension saving in the uk. *IFS Working Paper*.

Dahlquist, M., O. Setty, and R. Vestman (2018). On the asset allocation of a default pension fund. *The Journal of Finance 73*(4), 1893–1936.

De Nardi, M. and G. Fella (2017). Saving and wealth inequality. *Review of Economic Dynamics 26*, 280–300.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature 47*(2), 315–72.

DellaVigna, S. (2018). Structural behavioral economics. In *Handbook of Behavioral Economics: Foundations and Applications*.

DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017). Reference-dependent job search: Evidence from hungary. *The Quarterly Journal of Economics 132*(4), 1969–2018.

Dobrescu, L. I., X. Fan, H. Bateman, B. R. Newell, A. Ortmann, and S. Thorp (2018). Retirement savings: A tale of decisions and defaults. *The Economic Journal 128*(610), 1047–1094.

Franco, C. (2004). Unemployment benefits reduced by pensions and social security: A fact sheet.

Goda, G. S., M. R. Levy, C. F. Manchester, A. Sojourner, and J. Tasoff (2020). Who is a passive saver under opt-in and auto-enrollment? *Journal of Economic Behavior & Organization 173*, 301–321.

Goldin, J. and D. Reck (2020). Optimal defaults with normative ambiguity. *Review of Economics and Statistics*.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic dynamics 12*(1), 58–79.

Haggag, K. and G. Paci (2014). Default tips. *American Economic Journal: Applied Economics 6*(3), 1–19.

Havránek, T. (2015). Measuring intertemporal substitution: The importance of method choices and selective reporting. *Journal of the European Economic Association 13*(6), 1180–1204.

Illanes, G. (2016). Switching costs in pension plan choice.

Kaplan, G. and G. L. Violante (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica 82*(4), 1199–1239.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Laibson, D., P. Maxted, A. Repetto, and J. Tobacman (2017). Estimating discount functions with consumption choices over the lifecycle.

Low, H., C. Meghir, and L. Pistaferri (2010). Wage risk and employment risk over the life cycle. *American Economic Review 100*(4), 1432–67.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy 125*(2), 431–477.

Madrian, B. C. and D. F. Shea (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly journal of economics 116*(4), 1149–1187.

Masatlioglu, Y. and E. A. Ok (2005). Rational choice with status quo bias. *Journal of Economic Theory 121*, 1–29.

Menzio, G., I. A. Telyukova, and L. Visschers (2016). Directed search over the life cycle. *Review of Economic Dynamics 19*, 38–62.

O'Dea, C. (2018). Insurance, efficiency and the design of public pensions.

O'Donoghue, T. and M. Rabin (1998). Procrastination in preparing for retirement. In *Behavioral Dimensions of Retirement Economics. Washington DC and.* Citeseer.

Ramsey, F. P. (1928). A mathematical theory of saving. *The economic journal 38*(152), 543–559.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics 117*(3), 1039–1073.

Samuelson, W. and R. Zeckhauser (1988). Status quo bias in decision making. *Journal of risk and uncertainty 1*(1), 7–59.

Sleet, C. and Ş. Yeltekin (2006). Credibility and endogenous societal discounting. *Review of Economic Dynamics 9*(3), 410–437.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters 20*(2), 177–181.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American economic review 93*(2), 175–179.

Zhong, M. (2020). Optimal default retirement saving policies: Theory and evidence from oregonsaves.