

Present Bias Causes and Then Dissipates Auto-Enrollment Savings Effects

By JOHN BESHEARS, JAMES CHOI, DAVID LAIBSON, AND PETER MAXTED*

Auto features in defined contribution pension plans, like auto-enrollment, have a large short-run effect on participation and contribution rates (Madrian and Shea 2001). However, a growing literature finds that auto features may have only a modest average impact in the long run.¹

In this paper, we show that if households have present bias,² auto-enrollment is predicted to have a positive causal impact on savings in the defined contribution savings plan *at the household's current employer*. However, with sufficient present bias, the positive effect on current/proximate savings will be partially or even fully dissipated before retirement.

We provide a model that highlights two key channels that drive these opposing effects. First, present bias engenders procrastination, which leads households to stick with auto-enrollment defaults (Carroll et al. 1997). However, present bias also engenders over-consumption. Separation from an employer generates a rollover of 401(k) balances to an IRA account. Rollover IRA accounts are more liquid than 401(k) accounts. Distributions from an IRA are allowed for any reason, can fully deplete the IRA, and are penalty-free for some categories of spending³ or if the beneficiary is over age 59½. Even when the standard 10% early withdrawal penalty does apply to an IRA distribution, households with suf-

ficient present bias will be willing to partially or fully deplete these accounts before retirement.⁴

Without present bias (i.e., when the discount function is exponential), our illustrative/toy model generates no auto-enrollment savings effects and no follow-on leakage effects. Auto-enrollment doesn't affect equilibrium outcomes in our exponential model because we assume, for illustrative purposes, vanishingly small transactions costs for changing one's contribution rate in a 401(k) plan.

In contrast, we show that present-biased agents may be whipsawed by auto-enrollment. They follow their employer's default while still employed and then spend some or all of the new savings after they separate from their employer and the savings becomes more liquid.

I. Model

We present a model that illustrates why present bias makes auto-enrollment effective at one's current employer but *also* partially or fully reverses these effects in the long run due to pre-retirement distributions from the IRA. The model is stylized to highlight key mechanisms.

A. Preferences and Beliefs

We set the model in continuous time and assume that households have present-biased preferences. Specifically, the current self has a discount function that puts full weight on utils experienced in the immediate present and weight $\beta e^{-\rho\tau}$ on utils experienced at delay $\tau > 0$. When $0 < \beta < 1$, this is the instantaneous gratification model (Harris and Laibson 2011), which provides a tractable way of capturing present bias in continuous time and closely approximates the quantitative effects of present bias in discrete time (Laibson and Maxted 2022). In this notation, β is the present bias parameter and

* The authors thank Jessica Schultz for research assistance, Laura Quinby for comments, and the Pershing Square Fund for Research on the Foundations of Human Behavior for funding. Beshears: Harvard Business School, Baker Library, Soldiers Field, Boston, MA 02163, jbeshears@hbs.edu. Choi: Yale School of Management, 165 Whitney Ave., P.O. Box 208200, New Haven, CT 06520-8200, james.choi@yale.edu. Laibson: Harvard University Department of Economics, Cambridge, MA 02138, dlaibson@harvard.edu. Maxted: Haas School of Business, University of California, Berkeley, 545 Student Services Building 1900, Berkeley, CA 94720, maxted@haas.berkeley.edu.

¹Choi et al. (2004), Choukhmane (2021), Beshears, Choi, and Laibson (2021)

²See Strotz (1955) and Laibson (1997) for intra-personal, and Phelps and Pollak (1968) for intergenerational, present bias.

³Examples of penalty-free withdrawals include educational expenses, home buying/building, and medical expenditures.

⁴See Argento et al. (2013) for evidence on leakage.

ρ is the long-run discount rate. For the special case $\beta = 1$, this discount function is exponential.

We will study the case of fully naive beliefs (Strotz 1955, Akerlof 1991, O’Donoghue and Rabin 1999), where the household believes future selves will choose as if they have discount functions with $\beta = 1$. If we were to assume partial sophistication, equilibrium behavior would change in some ways, but the general conclusion would remain the same: auto-enrollment would have a large short-run effect and a modest long-run effect on wealth formation. With complete sophistication about future β , auto-enrollment would have little effect at any horizon.

To simplify exposition, we describe the case in which flow utility is the natural log of consumption, but the analysis generalizes to any constant relative risk aversion utility function.

B. Income and the Household Balance Sheet

The household cycles through N employment spells that each have duration T , with no gaps between spells, implying a total working life of NT years. During employment, the household receives income flow y . After working life, the household begins retirement, which is assumed to be infinitely long. During retirement, the household receives exogenous ‘labor’ income flow y_R (i.e., social security and/or some other defined benefit pension). As we explain next, the household can supplement retirement income with distributions from savings.

The household has access to three kinds of accounts, each of which has real return r :

1. A fully illiquid *workplace retirement account* with balance $z(t)$, which is modeled to capture the key properties of a 401(k) or similar employer-based defined contribution account. Contributions to this account are voluntary, and employee contributions up to m fraction of income earn matching contributions at a rate M .⁵ We will study two cases: (i) employees are automatically enrolled at savings rate $s = s_D$ when they start a new employment spell, and (ii) employees have a default savings rate $s = 0$ and need to opt into this account if they want a non-zero savings rate. We assume that workers must pay a lumpy effort cost E to change their savings

⁵In the United States, the most common DC configuration in 2020 is $m = 0.06$ and $M = 0.5$. See Vanguard (2021, p. 21).

rate away from the default. To highlight the role of lumpy utility costs in models of present bias, we follow Laibson, Maxted and Moll (2021) and study the case where E is vanishingly small but strictly positive.⁶ The dynamic budget constraint for z is given by $\dot{z} = sy + rz$ during employment spells. At the end of each employment spell, z drops discretely back to zero because we assume that the funds from account z are rolled over to an individual retirement account, which we discuss next.

2. A partially illiquid *individual retirement account* (IRA) that automatically receives rollovers from workplace retirement accounts after each employment spell ends.⁷ The IRA has an early withdrawal penalty of $p > 0$. Specifically, if funds are distributed from this account before retirement, the agent receives $1 - p$ dollars for every dollar distributed. During retirement, there is no early withdrawal penalty from the IRA. We use $w(t)$ to denote the balance in this account. Whenever an employee leaves an employer, w jumps up discretely by amount z . Excluding these jump events, the dynamic budget constraint for w is given by $\dot{w} = -d + rw$, where d represents the flow of distributions.

3. A liquid *savings account* with a weakly positive balance (i.e., no borrowing). If the agent begins economic life with no funds in this account, for the cases that we discuss below the liquid savings account is weakly dominated on the equilibrium path.⁸

In summary, WLOG we assume the household chooses to do its accumulation and decumulation exclusively using the two retirement accounts: the defined contribution account, $z(t)$, and the rollover IRA, $w(t)$. On the equilibrium path, total financial wealth is given by the sum $z(t) + w(t)$. During working life, flow consumption is $c(t) = y(1 - s) + d(t)(1 - p)$; during re-

⁶Because our model has no uncertainty, the existence of the match and a vanishingly small enrollment cost makes account z economically dominate outside savings accounts.

⁷In 2020 slightly more than half of the 401(k) accounts of employees who separated were either distributed as cash payouts at separation or rolled over within that calendar year (Vanguard 2021 p. 109). If a 401(k) balance at separation is under \$1,000 an employer can compel a cash distribution. If a 401(k) balance is between \$1,000 and \$5,000 an employer can compel an IRA rollover.

⁸In richer economic settings, a liquid savings account would serve as a source of high frequency liquidity.

tiement, flow consumption is $c(t) = y_R + d(t)$.

For expositional simplicity we assume: the interest rate equals the long-run discount rate ($r = \rho$) and $u'(y) > \beta u'(y_R)$.⁹

II. Equilibrium for an Exponential Household

We first study the case $\beta = 1$. We begin by assuming that the plan has an opt-in structure. To simplify exposition, assume that exogenous retirement income, y_R , and the match threshold, m , are sufficiently small that it is optimal for the exponential agent to save during working life at a rate greater than the match threshold m .¹⁰ In equilibrium, the Euler Equation with In utility implies $\dot{c} = (r - \rho) = 0$. Because the exponential agent never pays an early withdrawal penalty, equilibrium consumption is

$$c = \rho \left[\int_0^{NT} (1 + mM)y e^{-r\tau} d\tau + \int_{NT}^{\infty} y_R e^{-r\tau} d\tau \right].$$

Consumption does not change between working life and retirement.

To generate a quantitative benchmark, calibrate $\rho = r = 0.01$, $N = 10$, $T = 5$, $y = 1$, $m = 0.06$, $M = 0.50$, and $y_R = \frac{2}{3}y$. In equilibrium, the constant consumption level will be $0.81 \times y$ and the stock of wealth at retirement will be $14 \times y$ (see Appendix for calculations).

Because the effort cost of changing one's savings rate is arbitrarily small, this equilibrium is not affected by the default savings rate in the workplace retirement account; that is, automatic enrollment has no impact when $\beta = 1$.

III. Equilibrium with Present Bias

We start with an *opt-in* regime for the employer's retirement savings plan. With naive beliefs, the household predicts that it will join the retirement savings plan in the immediate future, but it fails to implement this plan because (lumpy) effort cost E is paid in the immediate present and it is preferable for a household to postpone this effort cost to the immediate future when it will be discounted with factor β . The

cost of postponement is perceived to be proportional to dt and the benefit of postponement is perceived to be a stock of utils: $(1 - \beta)e$. Accordingly, the (naive) household keeps postponing and never enrolls in the retirement savings plan (and also never saves in the liquid asset because $u'(y) > \beta u'(y_R)$), so the household retires with no savings. Consumption is equal to y during working life and y_R in retirement.

Under *auto-enrollment* (i.e., opt-out), the household remains at the auto-enrollment default savings rate for the same reason that the household remains at the 0 savings rate in the opt-in regime; the household expects that it will optimize in the immediate future but never actually does so. The differential equation for workplace retirement savings is $\dot{z} = s_D y + rz$. Before retirement, the household has flow consumption $c(t) = (1 - s_D)y$ when the IRA has no balance: $w(t) = 0$.

When $w(t) > 0$, the household will use the following Euler Equation *system* to choose consumption. In this system, notional consumption levels \hat{c} , c^* , and c^{**} are used as inputs for the calculation of equilibrium consumption, c .

The first equation establishes the level of permanent income, which is what the household mistakenly expects future selves to consume:

$$\hat{c}(t) = \rho [h(t) + z(t) + w(t)],$$

where $h(t)$ is the time- t present value of labor income (y), matching contributions, and exogenous retirement income (y_R). Note that $\hat{c}(t) \geq y_R$ (see Appendix).

The second equation establishes the relationship between permanent income (\hat{c}) and actual consumption *if* consumption is being funded (at the margin) by early withdrawals from the IRA:

$$(1 - p)u'(c^*(t)) = \beta u'(\hat{c}(t)).$$

If $c^*(t) < y(1 - s_D)$, then a different Euler Equation emerges because the household does not need to dissave from the IRA to consume at least flow $c^*(t)$. Accordingly, the penalty term drops from the Euler Equation:

$$u'(c^{**}(t)) \geq \beta u'(\hat{c}(t)).$$

Because we assumed that $u'(y) > \beta u'(y_R)$, it will always be the case that $c^{**}(t)$ is equal to

⁹With these assumptions, the decline in income at retirement is modest enough that a present-biased household won't save for retirement in a frictionless economy, but the household may save if their employer adopts a DC pension.

¹⁰See the Appendix for the necessary inequality.

the upper bound for consumption when distributions are not being taken from the IRA rollover: $y(1 - s_D)$. It follows that in equilibrium,

$$c(t) = \max\{c^*(t), y(1 - s_D)\}.$$

These four equations characterize consumption when the household has a non-zero IRA balance. With sufficiently low values of β the consumption path has a saw-toothed pattern, with consumption temporarily rising after each employment separation, when rollovers from relatively illiquid 401(k)s to relatively liquid IRAs generate transitory surges in consumption.

Figure 1 illustrates the consumption path (using the earlier calibration) for a present-biased household with $\beta = 0.65$. In this toy model, the household experiences a consumption surge after each of the first six (five-year) employment spells; the first five surges are large enough to fully deplete the household's IRA rollover account (until the next separation repopulates the rollover account). Starting with the employment spell that begins at age 50, the household avoids decumulating all of its IRA rollover wealth during each employment spell; the household recognizes that its stock of wealth is lower than it had anticipated it would be, so the retirement savings motive is strengthening.¹¹

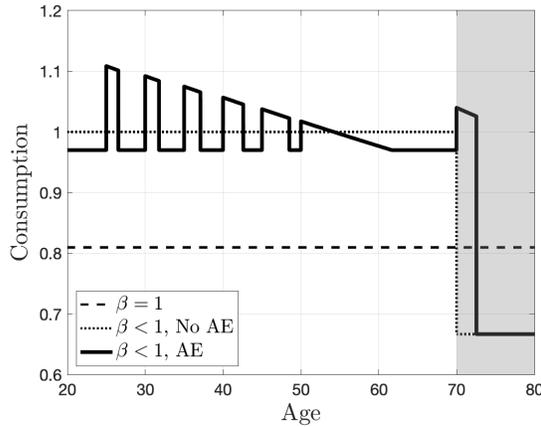


FIGURE 1: CONSUMPTION IN EXPONENTIAL CASE (DASHED FLAT LINE), PRESENT BIAS WITHOUT AUTO-ENROLLMENT (DOTTED LINE WITH AGE-70 CLIFF), AND PRESENT BIAS WITH AUTO-ENROLLMENT (SAWTOOTH).

¹¹The jump up at retirement (age 70) reflects two forces: (i) the release of funds from the last employer's 401(k) plan, and (ii) the elimination of the early withdrawal penalty. (In the U.S., the early withdrawal penalty actually ends at age $59\frac{1}{2}$.)

A. Sufficient condition for complete leakage

Complete leakage is a draw down of all IRA balances before the end of the household's final employment spell. In the appendix we derive a sufficient condition for complete leakage. We present an approximation, which matches to two significant digits for our calibration:

$$\beta \leq \left(\frac{1-p}{1+mM} \right) \left(\frac{y_R}{y} \right).$$

The model predicts that households with a β value below this threshold ($\beta = 0.58$ in our calibration) will *completely* dissave their IRA rollover balances from their first $N - 1$ employment spells. The only wealth that has not been spent by the time the household reaches retirement is the (default-based) 401(k) savings generated during their final employment spell.

The higher the penalty for early withdrawal, the lower β must be to generate penalty-based withdrawals from a rollover IRA. The higher the matching funds, mM , the lower the threshold for β , simply because more funds are in the IRA rollover (requiring a greater flow of pre-retirement distributions to achieve full depletion). The higher the labor-income replacement ratio, $\frac{y_R}{y}$, the higher the β threshold, because the motive for retirement saving is lower, thereby increasing willingness to take distributions.

B. Sufficient condition for some leakage

In the appendix we also provide a sufficient condition for partial leakage. If

$$\beta < \left(\frac{1-p}{1-s_D} \right) \left(\frac{\hat{c}(t)}{y} \right),$$

then the household will spend at least some of its IRA wealth early in life. This sufficient condition is less stringent (admitting higher values of β than the first bound) because early in life, households believe that they are going to be successful savers going forward and accordingly are more willing to spend IRA savings (anticipating that they won't have a pressing need for those savings later in life).

The higher the auto-enrollment default savings rate, s_D , the higher the threshold value of β because a high default savings rate lowers resources for immediate consumption and elevates

marginal utility, thereby increasing the motive for spending from a rollover IRA.

In the United States, the penalty for early withdrawals is $p = 0.10$, and the modal default savings rate is $s_D = 0.03$.¹² Using our calibration with t close to 0, the bound for β is 0.75. In other words, in our model, households with $\beta < 0.75$ will leak at least some of their defined contribution accumulation before retirement.¹³

IV. Conclusion

In a model with exponential discounting and small transactions costs, the introduction of auto-enrollment does not cause defined contribution savings behavior to change. However, in a model with present bias and naive beliefs, auto-enrollment substantially changes equilibrium behavior, causing employees to stick with the default savings rate. Because of present bias, households may leak savings from their rollover IRA. For households with sufficiently low (but empirically plausible) β values, leakage will cause a complete draw-down of their IRA rollover accounts before they reach retirement. Specifically, they will enter retirement only with defined contribution wealth accumulated at their last employer.

This highly stylized analysis illustrates the potentially adverse role that IRA liquidity may have in the U.S. savings system. If defined contributions were fully illiquid (like the systems in many other countries¹⁴), then retirement savings could not leak out before retirement.¹⁵ In the stylized setting of our model with no uncertainty, retirement account illiquidity would not reduce households' willingness to make contributions in the first place because they don't expect to draw on those assets before retirement. Accordingly, the first-best could be obtained simply by making IRAs fully illiquid with compulsory annuitization. In more realistic models with spending shocks (e.g., unexpected medical costs), liquidity could play a constructive role.

¹²Data from 2020 reported by Vanguard (2021 p. 28).

¹³Our model assumes $u'(y) \geq \beta u'(y_R)$, which implies $\beta \leq \frac{2}{3}$, because we calibrate $y_R = \frac{2}{3}$. Hence, in our calibration all households leak some of their IRA wealth. This would not be the case with a sufficiently high penalty for early withdrawals.

¹⁴See Beshears et al. (2015).

¹⁵However, see Maxted (2020) for a framework where households never hit liquidity constraints because they can always access an increasingly costly borrowing margin.

V. References

- Argento, Robert, Victoria Bryant, John Sabelhaus. 2013. "Early Withdrawals from Retirement Accounts during the Great Recession." *Contemporary Economic Policy*. 33. 10.2139.
- Akerlof, George A. 1991. "Procrastination and Obedience." *American Economic Review* 81:2, 1-19.
- Beshears, John, James J. Choi, Joshua Hurwitz, David Laibson, and Brigitte C. Madrian. 2015. "Liquidity in Retirement Savings Systems: An International Comparison." *AEA Papers and Proceedings* 105: 420-25.
- Beshears, John, James J. Choi, and David Laibson. 2021. "Do Automatic Savings Policies Actually Increase Savings?" Mimeo.
- Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124: 1639-74.
- Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2004. "For Better or For Worse: Default Effects and 401(k) Savings Behavior." In *Perspectives on the Economics of Aging*, edited by David A. Wise, 81-121. Chicago: University of Chicago Press.
- Choukhmane, Taha. 2021. "Default Options and Retirement Saving Dynamics." Mimeo.
- Harris, Christopher, and David Laibson. 2013. "Instantaneous Gratification." *Quarterly Journal of Economics* 128: 205-48.
- Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112: 443-78.
- Laibson, David, Peter Maxted, and Benjamin Moll. 2021. "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy." NBER Working Paper 29094
- Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116: 1149-87.
- Maxted, Peter. 2020. "Present Bias in Consumption-Saving Models: A Tractable Continuous-Time Approach." Mimeo.
- O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89: 103-24.

Phelps, Edmund S., and Robert A. Pollak. 1968. "On Second-Best National Saving and Game-Equilibrium Growth." *Review of Economic Studies* 35: 185-99.

Strotz, Robert Henry. 1955. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* 23: 165-180.

Vanguard. 2021. "How America Saves 2021." institutional.vanguard.com/how-america-saves (accessed December 26, 2021).