Windfall Gains and Stock Market Participation*

Joseph Briggs†
Federal Reserve Board of Governors

David Cesarini
New York University, IFN & NBER

Erik Lindqvist
Stockholm School of Economics & IFN

Robert Östling
Institute for International Economic Studies

We estimate the causal effect of wealth on equity market participation using data on Swedish lottery players. A $150,000 windfall gain increases stock market participation by 12 percentage points among pre-lottery nonparticipants, but has no discernible effect on stock owners. A life-cycle model overpredicts rates of entry and requires implausibly large entry costs to match our quasi-experimental estimates. Effects are larger among households with more education, greater cognitive ability, and that won in years following positive equity returns, suggesting nonstandard beliefs or information costs contribute to the limited effect of wealth on stock market participation.

*This paper is part of a project hosted by the Research Institute of Industrial Economics (IFN). We are grateful to IFN Director Magnus Henrekson for his strong support of the project and to Marta Benkestock for superb administrative assistance. For helpful comments, we thank Steffen Andersen, Dan Benjamin, Claudio Campanale, David Laibson, Annette Vissing-Jørgensen, Enrichetta Ravina, Paolo Sodini, and Roine Vestman, as well as seminar participants at NYU Stern, IFN, Institute for International Economic Studies, London Business School, Copenhagen Business School, the Federal Reserve Bank of New York, the Federal Reserve Board of Governors, University of Southern California, the 2015 NBER SI (Household Finance), and the 2015 SED. We also gratefully acknowledge financial support from the Alfred P. Sloan Foundation through the NBER Household Finance small grant program (22-2382-13-1-33-003), the NSF (1326635), the Swedish Council for Working Life and Social Research (2011-1437), the Swedish Research Council (B0213903), Riksbankens Jubileumsfond (P15-0615:1), and Handelsbanken's Research Foundations (P2011:0032:1). The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Board of Governors.
JEL No: D1,D14,D91, G02,G11, G12

†Corresponding author: Joseph Briggs. Board of Governors of the Federal Reserve System, Constitution Ave & 20th St NW, Washington, DC 20551. 202-245-4214. joseph.s.briggs@frb.gov.

1 Introduction

Canonical life-cycle models of consumption and savings predict that all individuals should invest a positive fraction of their wealth in equities (Samuelson (1969); Merton (1971)). However, a sizable fraction of households in developed countries do not own equity (Guiso, Haliassos and Jappelli (2002)). A large literature in household finance formulates and tests hypotheses about the causes of this "nonparticipation puzzle" (Haliassos and Bertaut (1995); Vissing-Jørgensen (2003); Campbell (2006); Guiso and Sodini (2013)), as insights into the causes of equity market nonparticipation may guide efforts to promote efficient financial decision-making (Campbell (2006)).

Limited stock market participation is often analyzed in models where agents weigh the benefits of owning equities against its costs (Mulligan and Sala-i Martin (2000); Vissing-Jørgensen (2003); Paiella (2007); Attanasio and Paiella (2011)). Early work in this area by Vissing-Jørgensen (2003) posited a simple framework with two types of fixed costs: per-period participation costs and a one-time entry cost. Since the gains from participation are increasing in wealth, whereas the costs are assumed fixed, these models provide a simple interpretation of the positive correlation between wealth and stock market participation (Mankiw and Zeldes (1991); Poterba and Samwick (2003); Campbell (2006)). This framework has subsequently been extended by a large structural literature which models the lifecycle saving and portfolio decisions of households (e.g., Gomes and Michaelides (2005); Cocco (2005); Alan (2006); Khorunzhina (2013); Fagereng, Gottlieb and Guiso (2017a)) to account for housing (Cocco (2005); Flavin and Yamashita (2011); Vestman (2013)), outstanding debt (Davis, Kubler and Willen (2006); Becker and Shabani (2010)), private business equity (Heaton and Lucas (2000a)) and stochastic labor income (Viceira (2001)).

Although such models of equity market participation make precise, quantitative predictions about the effect of a windfall gain on risk-taking behavior, credibly testing these predictions is difficult. In this paper, we estimate the causal effect of lottery prizes on participation in equity markets by exploiting the randomized assignment of wealth in three Swedish samples of lottery players who have been matched to administrative records with high-quality information about financial portfolios. The sample has a number of desirable characteristics. First, we observe the factors conditional on which lottery wealth is randomly assigned (e.g. number of tickets owned). Second, because the size of the prize pool is over 500 million USD, our study has excellent power to detect even modest effects of wealth on participation over various time horizons. Third, the prizes won by the players in our sample vary in magnitude, allowing us to explore and characterize nonlinear effects of wealth. Finally, because our lottery and financial data are drawn from administrative records, our sample is virtually free from attrition.

Our first contribution is thus to provide credible and precise estimates of the effect of large wealth shocks on stock market participation. The relationship between wealth and participation is usually estimated using observational data (Brunnermeier and Nagel (2008); Calvet and Sodini (2014)) where, even applying the best methods, it is difficult to completely eliminate concerns about omitted variables and simultaneity. Recognizing these difficulties, Carroll (2002) notes that the "ideal experiment to answer the causality question would be to exogenously dump a large amount of wealth on a random sample of households and examine the effect... on their risk-taking behavior" (p. 19). Our research design closely approximates Carroll's ideal experiment: prizes are randomly assigned conditional on factors that we can observe, and the identifying variation primarily comes from large wealth shocks. We find that each 1M SEK (approximately 150K USD) received increases the probability of stock ownership in post-lottery years by 4 percentage points.[1] This effect is driven almost entirely by an immediate and seemingly permanent 12 percentage point effect among households that did not participate in equity markets prior to winning the lottery.

As noted in Kahn and Whited (2016), causal estimates alone provide limited insight into the economic forces that influence participation decisions. Our second contribution, therefore, is to use a structural life-cycle model to interpret the lottery estimates. We first show that our model predicts far larger effects than the reduced-form estimates when using parameters estimated from non-experimental data. We then estimate model parameters implied by our causal estimates, most notably the implied costs of equity market entry and participation. Accounting for participation responses to lottery wins requires entry costs that are extremely large: the average entry cost for pre-lottery equity market nonparticipants is 31K USD, or approximately 10 times larger than the average cost estimated from non-experimental data. Our structural estimation thus demonstrates the challenge our reduced-form causal estimates pose to standard models of stock market participation.

As a final contribution, we examine the credibility of alternative explanations for stock market nonparticipation. We find limited evidence that pathways related to other economic incentives, including debt reduction and real estate investment, account for the relatively small causal effects and large entry cost estimates. We similarly find little evidence that alternative preference

[1] All monetary variables presented in this paper are reported in year-2010 prices. When converting to USD, we use the Dec. 31, 2010 exchange rate of 6.72 SEK/USD.

specifications – including status quo bias, loss aversion, and present-biased preferences – account for our empirical results. However, we do find substantial evidence in favor of pathways related to information and beliefs. For example, treatment effects are larger for more educated and more cognitively able lottery players. Furthermore, larger effects among households that won after years of positive equity returns or that experienced above median equity returns between ages 18-25 suggest the importance of non-standard belief-formation processes (Choi and Mertens (2013); Greenwood and Shleifer (2014)). The credibility of beliefs and information is bolstered by additional structural analyses. The difference between model predictions and empirical estimates decreases when we restrict the sample to highly educated, high cognitive ability households and when we account for subjective beliefs regarding equity returns. Overall, our results point strongly towards non-standard beliefs or information costs as important factors behind stock market nonparticipation.

Most closely related to our work is Andersen and Nielsen (2011) who make sophisticated use of Danish administrative data to study how inheritances caused by sudden deaths affect stock market participation. They compare stock market participation of such inheritors to individuals matched on age, sex, education, and earnings and wealth deciles. There are two key methodological differences between our reduced-form analyses and Andersen and Nielsen (2011). First, a bequest due to a sudden death is conceptually different from a windfall gain to lifetime wealth. Although unexpected inheritances clearly increase liquid wealth, the net impact on lifetime wealth is difficult to quantify (or perhaps even sign correctly) absent further assumptions on the parent's saving, investment and consumption decisions under the counterfactual scenario where the parent dies at an older age. Our study's estimates can be interpreted unambiguously as reflecting the causal impact of lottery-induced positive shocks to lifetime wealth.[2] Second, epidemiological literature have documented risk factors for the sudden deaths studied in Andersen and Nielsen (2011) (e.g., World Health Organization (2004)). A causual interpretation of their estimates hinges upon the assumption that risk factors influencing stock market participation are balanced across treatment and controls. We show in a series of stringent randomization checks that the wealth shocks we exploit are independent of pre-lottery characteristics, as expected under our identifying assumptions.

The paper is structured as follows. Section 2 describes the lottery and wealth data, our identification strategy, and addresses several issues regarding external validity that are often raised about

[2] Andersen and Nielsen (2011)'s treatment effect may also capture direct effects that sudden death may have on financial decision-making (e.g., the effects of grief on attitudes or economic behavior), as well as the (potentially heterogeneous) impacts of the different inherited asset classes. Some of these differences allow Andersen and Nielsen (2011) to explore interesting hypotheses for which our data are not suitable (e.g. investment inertia).

studies of lottery players. Section 3 reports reduced-form estimates of the effect of wealth on equity market participation, while Section 4 uses a structural life-cycle model to interpret the causal estimates. Section 5 presents a set of empirical and structural analyses to evaluate the credibility of alternative explanations of our results. Finally, Section 6 discusses our findings and concludes.

2 Data and Identification Strategy

Our analyses are conducted in a sample of lottery players who have been matched to administrative demographic and financial records using players' personal identification numbers (PINs).

2.1 Register Data

Our outcome variables are all derived from the Swedish Wealth Register, which contains high-quality information about the financial portfolios of all Swedes. The register was discontinued when Sweden abolished its wealth tax, but has annual year-end financial information for 1999-2007. This information includes total assets and debt, and relevant subcategories such as bank account balances, mutual funds, directly held stocks, bonds, money market funds, debt, and residential and commercial real estate. The data have proven valuable in household-finance research beginning with a landmark paper by Calvet, Campbell and Sodini (2007), who estimate that included variables account for approximately 84% of wealth in Sweden, but has a few notable data limitations. First, assets in private pension plans are not measured. Second, we do not observe the composition of capital insurance, an asset that can be composed of any financial assets. Third, Nekoei and Seim (2018) note that unlisted business equity is not measured in the wealth registry.

We supplement the portfolio data from the Wealth Register with basic demographic information available from Statistics Sweden. Our analyses are conducted at the household level, with a household defined as the observed winner and, if present, his or her spouse.[3] All our analyses are based on players aged 18 and above.

2.2 Lottery Data

Our identification strategy is to use the available data and knowledge about the institutional details of each of the lotteries to define cells within which the lottery wealth is randomly assigned. We then control for these cell fixed effects in our analyses, thus ensuring all identifying variation comes from players in the same cell. Because the exact construction of the cells varies across lotteries,

[3] Wealth of spouses of winning players increases by about 10% of the prize won following the lottery event, suggesting some joint control over assets. Table B.7 conducts analyses at the individual level as a robustness check.

we describe each lottery separately. For a more detailed description of the data, including how the original lottery data were preprocessed and quality-controlled, we refer the reader to Section 2 and the Online Appendix of Cesarini, Lindqvist, Östling and Wallace (2016). Unless otherwise noted, prizes are paid as a one-time lump-sum and all amounts are after tax. In this paper, all prize amounts (and other financial variables) are adjusted for inflation and expressed in year-2010 SEK and USD.

Kombi Kombi is a monthly subscription lottery whose proceeds are given to the Swedish Social Democratic Party, Sweden's main political party during the post-war era. Kombi provided us with a longitudinal data set with information about all draws conducted between 1998 and 2010. For each draw, the panel contains an entry per lottery participant, with information about the number of tickets held, any large prizes won, and the player's PIN.

In a given Kombi draw, each prize is awarded by randomly selecting a unique ticket. Two individuals who purchased the same number of tickets are equally likely to win a large prize. To construct the cells, each winning player is matched to (up to) 100 non-winning players with the same number of tickets in the month of the draw. To improve precision, we choose controls similar to the winner on sex and age whenever more than 100 matches are available. This matching procedure leaves a sample of 346 large prize-winners, matched to a total of 31,180 controls.

Triss Triss is a scratch-ticket lottery run since 1986 by Svenska Spel, the Swedish government-owned gambling company. Since 1994, Triss players can win an opportunity to participate in a TV show where they draw a prize by selecting a ticket from a shuffled stack. In our main analyses, we restrict the Triss sample to 3,404 players who won lump-sum prizes between 7.8K USD (52K SEK) and 909K USD (6.1M SEK). However, in Section 3.1 we compare estimates for lump-sum prize winners to a "Triss-monthly" sample of 476 players who received prizes paid in monthly installments over 10 to 25 years (see Appendix Table B.1 for descriptive statistics). We convert the installments to net present value to make them comparable to lump-sum prizes.

Svenska Spel supplied the basic demographic information (name, age, and address) of all TV show participants between 1994 and 2011, allowing us to identify 99% of participants. Svenska Spel also listed cases in which the player shared ownership of the ticket. Our analyses are based exclusively on the 93% of winners that did not indicate they shared ownership of the winning ticket. Our empirical strategy makes use of the fact that, conditional on the prize type (lump-sum or monthly installments), the nominal prize amount is random. Thus, two players are assigned to the same cell if they won the same type of prize, in the same year, and under the same prize plan.

PLS Prize-linked savings (PLS) accounts are savings accounts whose owners participate in regular lotteries with monetary prizes paid on top of (or sometimes in lieu of) interest payments. In Sweden, PLS accounts were subsidized by the government until 1985, at which point the government ceased subsidies but authorized banks to continue to offer PLS accounts. Two systems were put into place, one operated by savings banks and one by commercial banks and the state bank. The two systems were approximately equally popular and participation was widespread across broad strata of Swedish society, with every other Swede owning an account in the late 1980s.

The PLS sample was obtained by combining prize lists and monthly data on account balances from the PLS accounts maintained by commercial banks and the state bank. These data allow us to identify the account owner, account balance, and amount won in each draw. Overall, we were able to reliably identify the owner's PIN for 99% of prize-winning accounts. PLS account holders could win odds prizes or fixed prizes. The probability of winning either type of prize was proportional to the number of tickets associated with an account: account holders were assigned one lottery ticket per 100 SEK in account balance. Fixed prizes were prizes whose magnitude did not depend on the balance of the winning account. Odds prizes, on the other hand, were awarded as a multiple of the balance of the prize-winning account.

For fixed prize winners, our identification strategy exploits the fact that in the population of players who won exactly the same number of fixed prizes in a particular draw, the total amount is independent of the account balance (Imbens, Rubin and Sacerdote (2001); Hankins, Hoestra and Skiba (2011)). We therefore assign two individuals to the same cell if they won an identical number of fixed prizes in that draw. To construct odds prize cells, we match individuals who won exactly one odds-prize between 1989-1994 in a draw to individuals with a near-identical account balance who also won exactly one prize (odds or fixed) in the same draw. This matching procedure ensures that within a cell, the prize amount is independent of potential outcomes. In total, the sample includes 332,647 PLS prizes, of which 478 are larger than 1M SEK (150K USD).

2.3 Identification Strategy

Table 1 summarizes the previous section's discussion of how we construct the cell fixed effects in each of the three lotteries. Normalizing the time of the lottery to $s = 0$, our main estimating equation is given by,

$$Y_{i,s} = \beta_s L_{i,0} + \boldsymbol{X}_{i,0}\mathrm{M}_{\mathbf{s}} + \boldsymbol{Z}_{i,-1}\boldsymbol{\gamma}_s + \eta_{i,s}, \tag{1}$$

Table 1: Overview of Identification Strategy.

Lottery	Period	Prize Type	Cells
PLS	1989-2003	Fixed Prize	Draw $\times$ # Fixed Prizes
PLS	1989-1994	Odds Prize	Draw $\times$ Balance
Kombi	1994-2007	Fixed Prize	Draw $\times$ # Tickets
Triss Lump-sum	1994-2007	Fixed Prize	Year $\times$ Prize Plan
Triss Monthly	1997-2007	Fixed Prize	Year $\times$ Prize Plan

where i indexes households, $L_{i,0}$ denotes the prize size (in million SEK), $\boldsymbol{X}_{i,0}$ is a vector of cell fixed effects, and $\boldsymbol{Z}_{i,-1}$ is a vector of controls measured in the year before the lottery. We include the controls only to improve the precision of our estimates. Standard errors are clustered at the level of the player. The key identifying assumption needed for β_s to have a causal interpretation is that the prize amount won is independent of $\eta_{i,s}$ conditional on the cell fixed effects.

We estimate Equation 1 in our pooled sample and in the subsample of players who participated in draws conducted between 2000 and 2007. In what follows, we refer to these samples as the *all-year* and the *post-1999* samples. The post-1999 sample plays an important role in subsample analyses where we stratify players by their pre-lottery participation status, which is first observed in 1999. In the all-year sample regressions, the set of pre-lottery controls include age, sex, marital status, higher education, household size, household income, and Nordic born. In the post-1999 sample regressions, we additionally control for net wealth, gross debt, and an indicator for real estate ownership.

Prize Variation To get a better sense of the source of our identifying variation, Table 2 summarizes the distribution of prizes. The total value of the after-tax prize money disbursed to the winners in our samples is over 500M USD (3.4B SEK). Although most prizes are small, our lottery-based estimates are mostly informative about the effect of winning large sums of money. Most of the identifying variation in all three lotteries comes from within-cell comparisons of non-winners, or winners of small or moderate amounts, to large-prize winners. One way to see this is to consider the change in the total treatment variation when prizes of different sizes are dropped from the data.[4] Dropping the 308,948 prizes below 1.5K USD (10K SEK) in the all-year sample reduces the treat-

[4]Because we are controlling for the cell fixed effects, the treatment variation used to identify the effect of wealth is the total sum of squares of prizes demeaned at the level of the cell.

Table 2: Prize Distribution. Included are the pooled all-year and post-1999 samples, and their respective lottery subsamples. Prize amounts are in year-2010 USD and net of taxes.

Prize Amount	A. All-year				B. Post-1999			
(K USD)	Pooled	PLS	Kombi	Triss	Pooled	PLS	Kombi	Triss
$L_i = 0$	31,180	0	31,180	0	26,126	0	26,126	0
$L_i \leq 1.5K$	308,948	308,948	0	0	41,578	41,578	0	0
$1.5K < L_i \leq 15K$	22,082	21,097	0	985	734	368	0	366
$15K < L_i \leq 75K$	4,009	1,935	0	2,074	1,237	0	0	1,237
$75K < L_i \leq 150K$	346	189	0	157	89	0	0	89
$150K < L_i \leq 300K$	822	443	330	49	297	2	273	22
$300K < L_i$	190	35	16	139	78	0	16	62
N	367,577	332,647	31,526	3,404	70,139	41,948	26,415	1,776

ment variation by 1.4% while dropping the 1,012 prizes above 150K USD (1M SEK) reduces the treatment variation by 91.1%.[5] All three lotteries contribute substantial identifying variation to the all-year sample, but Kombi and Triss prizes jointly account for most identifying variation in the post-1999 sample.[6]

Testing for Random Assignment To test our key identifying assumption, we again normalize the time of lottery to $s = 0$ and run the following regression:

$$L_{i,0} = \boldsymbol{X}_{i,0}\boldsymbol{\Gamma}_0 + \boldsymbol{Z}_{i,-1}\boldsymbol{\rho}_{-1} + \epsilon_i. \tag{2}$$

Under the null hypothesis of conditional random assignment, the characteristics determined before the lottery ($\boldsymbol{Z}_{i,-1}$) should not predict the lottery outcome ($L_{i,0}$) conditional on the cell fixed effects ($\boldsymbol{X}_{i,0}$). We run these randomization tests in the pooled all-year and post-1999 samples, and for each lottery separately in the post-1999 sample. As expected, Table 3 shows that the lagged characteristics have no statistically significant predictive power in the specifications that include cell fixed effects. If they are omitted however (columns 2 and 4), the null hypothesis of random

[5]We keep the non-winners in the Kombi lottery in this exercise. Since almost all winners in the Kombi lottery won a prize close to 150K USD (1M SEK), removing the non-winners would take out all identifying variation from this lottery.

[6]In the all-year sample, Triss contributes 57% of the treatment variation, Kombi 14% and PLS 29%. The corresponding number in the post-1999 sample is 64% for Triss lottery, 35% for Kombi and less than 1% for PLS.

Table 3: Testing for Random Assignment. Results are obtained by estimating Equation 2 in our all-year sample, in the post-1999 sample, and in the post-1999 lottery subsamples.

	All-year		Post-1999				
	Pooled		Pooled		PLS	Kombi	Triss
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fixed Effects	Cells	None	Cells	None	Cells	Cells	Cells
Demographic Controls							
F-stat	0.80	9.92	1.13	8.41	0.69	1.41	1.34
p	0.61	0.00	0.33	0.00	0.72	0.22	0.21
Financial Controls							
F-stat			1.29	17.38	0.77	0.87	1.22
p			0.28	0.00	0.51	0.46	0.30
Demographic + Financial Controls							
F-stat			1.52	14.95	0.81	1.65	1.43
p			0.11	0.00	0.64	0.11	0.15
N	367,577	367,577	70,139	70,139	41,948	26,415	1,776

assignment is rejected.

2.4 Representativeness of the Lottery Sample

The main concern about the external validity of our sample is that individuals who play the lottery may not be representative of the population at large. To investigate representativeness, we compare the lottery samples, weighted by prize size, to randomly drawn population samples of adult Swedes matched on sex and age.

Columns 1 and 2 of Table 4 show that the demographic characteristics of our lottery players closely resemble those of the representative sample. Columns 3 and 4 compare the financial characteristics of members of the post-1999 sample to a matched population sample. The pooled lottery sample has slightly less wealth than the matched population sample, slightly more debt, and is slightly more likely to own real estate. Notably, the equity market participation rate (the main outcome in our study) in our pooled sample is 66%, close to the 63% participation rate in the matched population sample. Columns 5-7 provide the corresponding descriptive statistics for the post-1999 sample broken down by lottery. PLS participants, who are selected on bank account

Table 4: Representativeness of All-Year and Post-1999 Samples. This table compares our prize-weighted all-year and post-1999 samples to representative samples matched on sex and age. The summary statistics shown are all means and measured at $s = -1$. All variables except female, age, and Nordic born are measured at the household level. Financial variables are winsorized at the .5 and 99.5 percentiles.

	All-Year		**Post-1999**				
	Pooled	**Pop**	**Pooled**	**Pop**	**PLS**	**Kombi**	**Triss**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Demographic							
Female	0.50	0.50	0.52	0.52	0.58	0.44	0.56
Age (years)	56.6	56.6	56.2	56.2	62.9	61.7	51.9
Nordic Born	0.96	0.93	0.96	0.92	0.95	0.98	0.94
Household Members	0.38	0.41	0.43	0.42	0.24	0.22	0.59
Household Income	323	302	364	362	330	342	382
Married	0.56	0.57	0.52	0.54	0.52	0.48	0.54
College	0.23	0.24	0.24	0.31	0.27	0.22	0.26
Financial							
Net Wealth (K USD)			131	158	205	123	128
Gross Debt (K USD)			53	49	27	37	67
Home Owner			0.75	0.69	0.73	0.78	0.73
Equity Market Participant			0.66	0.63	0.74	0.69	0.63
N	367,577		70,139		41,948	26,415	1,776

ownership, have significantly more wealth than the representative sample.

Another way to gauge representativeness is to compare the cross-sectional relationships between stock market participation and household characteristics in our lottery samples to the relationships estimated in a representative sample. We conduct such a comparison by estimating the cross-sectional probit equation used by Calvet et al. (2007) in their study of a large representative sample of Swedes. To avoid including wealth variation that was induced by the lottery, we restrict the estimation sample to the post-1999 sample and use observations the year prior to the lottery. We then repeat this regression for the matched representative sample. Appendix Table B.2 shows that the results from these regressions are quite similar.

While the absence of large differences in pre-lottery financial and demographic characteristics between our lottery sample and the representative sample is reassuring, we of course cannot rule

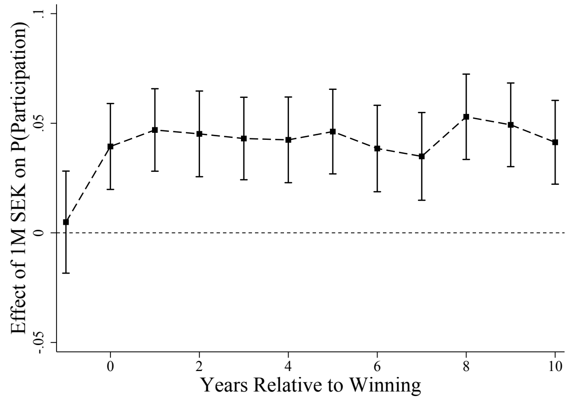


Figure 1: Effect of Wealth (1M SEK) on Participation Probability. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 in the all-year sample. See Appendix Table B.3 for the underlying estimates.

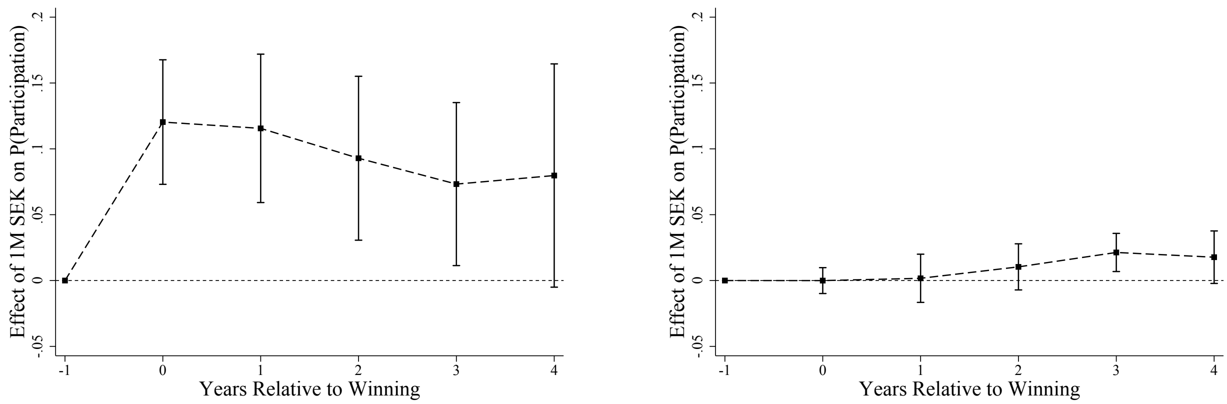
out the possibility that selection into lotteries based upon unobserved factors limits the external validity of our results.

3 Quasi-experimental Estimates

Our primary outcome variable is year-end participation, defined (as is standard in the literature) as an indicator variable equal to 1 for households that own stocks either directly or indirectly via mutual funds. Figure 1 presents the estimated coefficients for $s = -1, ..., 10$ from the all-year lottery sample. We estimate that each 150K USD (1M SEK) causes a near-immediate and permanent increase in the participation probability of around 3.9 percentage points. As expected, lottery wealth does not predict participation prior to the lottery. Effects are qualitatively similar but quantitatively smaller if we define participation more narrowly to only include directly owned stocks (Appendix Table B.3, Panel B).

We next investigate treatment effect heterogeneity with respect to equity market participation prior to the lottery. Figure 2 shows the estimated treatment effects on participation at $s = -1, ..., 4$ in the post-1999 sample stratified by pre-lottery participation status. Among pre-lottery nonparticipants, each 150K USD (1M SEK) increases participation by 12.0 percentage points at $s = 0$. The estimated treatment effect among nonparticipants is similar in the four years following the lottery, though less precisely estimated as we extend the time horizon.[7] In contrast, the estimated effect is

[7]There are two reasons why confidence intervals widen. First, participation is only observed during a nine-year



(a) Nonparticipants **(b) Participants**

Figure 2: Effect of Wealth (1M SEK) on Participation Probability by $s = -1$ Participation Status. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 in the post-1999 sample of nonparticipants (a) and participants (b). See Appendix Table B.4 for the underlying estimates.

small and mostly not statistically distinguishable from zero among pre-lottery participants. Hence, the aggregate effect of 3.9 percentage points we observe in the pooled sample appears to be driven nearly entirely by a positive effect on nonparticipants.

Because large prizes account for most of the identifying variation, our linear estimator assigns most weight to the marginal effect of wealth at modest to large levels of wealth. To test for non-linear effects, we replace the lottery-wealth variable in Equation 1 by indicator variables for five categories defined according by prize size. We then run regressions with the smallest prize category omitted.

Figure 3 presents the estimated coefficients for each of these categories, with coefficients marked at the mean prize size in each category. Relative to small prize winners (<1.5K USD, 10K SEK), a prize in the range 1.5 to 15K USD (10K-100K SEK) increases the participation probability of pre-lottery nonparticipants by .014. The corresponding estimates for winners of prizes in the 15 to 150K (100K-1M), 150 to 300K (1M-2M), and 300K+ (2M+) are .082, .177 and .399. Thus, the marginal effect (defined as the slope between points in Figure 3) is everywhere positive, but largest for winners of small prizes. Among pre-lottery participants, none of the prize-category coefficients are statistically distinguishable from zero.

period and we condition on prior participation status, so the sample size decreases with time horizon. Second, the predictive power of lagged financial and demographic characteristics falls with time, increasing the standard errors.

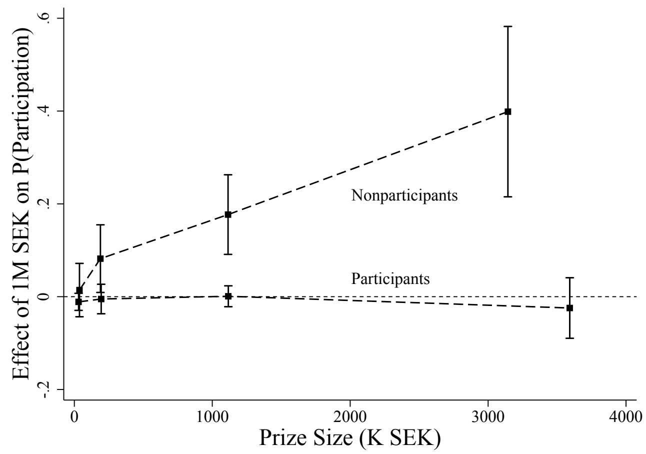


Figure 3: Effect of Wealth on Participation Probability by Prize Size. Coefficients are obtained by estimating Equation 1 in the post-1999 sample with the lottery wealth variable replaced by indicators for five mutually exclusive prize categories: 0 to 1.5K USD (0 to 10K SEK), 1.5 to 15K (10K to 100K), 15 to 150K (100K to 1M), 150 to 300K (1M to 2M), and 300K+ (2M+). Coefficient estimates and the 95% confidence bands are plotted at the mean prize in each category. See Appendix Table B.5 for the underlying estimates.

We conducted a number of sensitivity checks to explore the robustness of our $s = 0$ results. The results from these analyses are summarized in Appendix Table B.7. The estimated effect of lottery wealth on participation is similar across the Triss and Kombi lotteries.[8] Because selection into each of these two lotteries is markedly different, similarity in results across these subsamples strengthens confidence in the external validity of our results. In addition, our results are robust to excluding spousal equity ownership from our definition of participation. We also find that marginal effects from a probit estimator are quite similar to the OLS estimates reported in the main text.

Because capital insurance likely entails some equity exposure, we expand the definition of equity market participation to include ownership of capital insurance and find the effect of each 150K USD (1M SEK) among pre-lottery nonparticipants increases from 12 to 15 percentage points. Similarly, retail structured products in Sweden likely entail some equity exposure. In Section 5.2 we consider structured products and examine what inferences can be made from entry into this market. Finally, Appendix Table B.7 shows our main results are robust to dropping small (<7.5K USD, <50K SEK) prizes, but increase slightly when we drop large (>225K USD, >1.5M SEK) prizes.

[8] We exclude PLS from this comparison because, as noted in Section 2.3, PLS contributes little identifying variation to the post-1999 sample we focus on here.

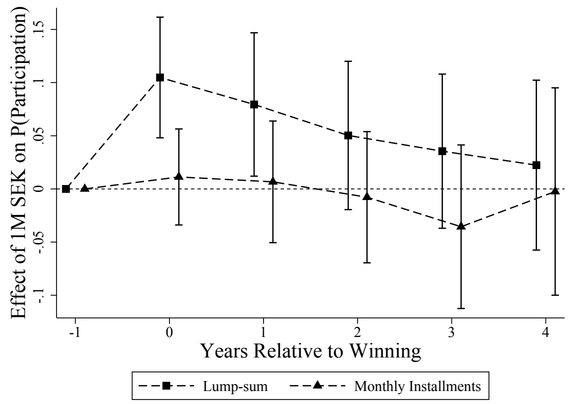


Figure 4: Effect of Wealth (1M SEK) on Participation by Payment Form. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 in the post-1999 sample of Triss winners stratified by payment plan type for nonparticipants. See Appendix Table B.6 for the underlying estimates.

The latter effect is expected in light of the decreasing marginal effect of lottery wins documented in Figure 3.

3.1 Type of Prize Payment: Lump Sum vs. Monthly Installments

The finding that even for nonparticipating households that won 300K+ USD (2M+ SEK) a majority do not buy stocks, suggest that many pre-lottery nonparticipants face a large hurdle to entering equity markets. This large disincentive could reflect either an initial entry cost or continuously paid participation costs. To help distinguish between these explanations, we exploit the "Triss-monthly" subsample that received monthly installments instead of a lump-sum prize. If up-front costs determine stock market participation and capital markets are imperfect, a liquid lump-sum prize would result in a larger effect on participation than illiquid monthly installments.

Figure 4 shows that among Triss players who did not own equities prior to the lottery, each 150K USD (1M SEK) in net present value paid in monthly installments had no effect on stock market participation probability the year of the lottery event. In contrast, each 150K USD (1M SEK) paid as a lump sum increases stock market entry at $s = 0$ by 10.5 percentage points. Furthermore, in subsequent years we find no evidence that prizes paid in monthly installments affected stock market participation, while the estimated effect of lump-sum prizes is positive (though not always statistically significant) at all horizons.[9] These differences by payment plan suggest that

[9] Appendix Table B.6 shows that – consistent with Figure 2 Panel (b) – there is no effect among pre-lottery equity market participants for either monthly installments or lump-sum prizes.

up-front costs are more likely to disincentive participation than continued costs of participation.[10]

In classical models with complete markets (e.g., Samuelson (1969)), participation and entry costs are equivalent and the household problem can be simplifed to a static setting. However, stark differences in the effects on entry by payment plan suggest that a simplified model (e.g., Vissing-Jørgensen (2003)) is insufficient to identify the structure of participation disincentives. Correct inference instead requires application of an appropriate economic framework, which we turn our attention to in the next section.

4 Structural Analysis

Previous structural work (e.g., Gomes and Michaelides (2005); Cocco (2005); Alan (2006); Khorunzhina (2013); Cooper and Zhu (2016); Fagereng et al. (2017a)) have shown that modest costs of entry and/or participation – which proxy for the totality of time costs, financial costs, and behavioral disincentives – are sufficient to disincentivize low-wealth households from purchasing equity and match observed participation patterns. In this section, we estimate a structural model to analyze whether this conclusion holds up also in our quasi-experimental data.

4.1 Model Specification

Each period, an age t agent chooses how much to consume C_t and save A_t – a fraction α_t of which is invested in equities – given their normalized cash on hand X_t, prior equity market participation status I_t, permanent income P_t, and lottery prizes L_t.

Demographics Each agent in our model is a single household with a fixed marital status $m \in \{0, 1\}$. Households fall into one of three education groups: high school education ($e = 0$); some post-secondary education ($e = 1$), and college degree or higher ($e = 2$). Life lengths are stochastic and finite – households survive from age t to $t + 1$ with probability s_t, and die with certainty at age $T = 100$ if they survive to that age.

[10] One complicating factor when comparing Triss-Lumpsum and Triss-Monthly is that the support of the prize distribution in the two lotteries differ (50,000 to 6M SEK in Triss-Lumpsum, a net present value of 1.1 to 10.5M SEK in Triss-Monthly). We therefore exclude Triss-Monthly prizes above 6M SEK in our analyses. Unshown analyses confirm the difference between Triss-Monthly and Triss-Lumpsum is robust to an alternative estimation strategy that uses the panel dimension of the data and compares winners before and after the lottery.

Preferences Epstein-Zin preferences (Epstein and Zin (1991)) are denoted

$$V_t = \left\{ (1-\beta s_t) C_t^{1-1/\psi} + \beta \mathbb{E}\left[s_t V_{t+1}^{1-\rho} + (1-s_t) b (X_{t+1})^{1-1/\psi} \right]^{\frac{1-1/\psi}{1-\rho}} \right\}^{\frac{1}{1-1/\psi}},$$

where C_t is consumption, β is the time discount factor, ρ is risk aversion, ψ is the intertemporal elasticity of substitution, and b is a bequest multiplier.

Income Each year alive, agents receive labor income Y_t. Before retirement, income is risky and follows the standard specification

$$\begin{aligned} Y_t &= exp(f(t,m,e)) P_t U_t \\ P_t &= P_{t-1} N_t, \end{aligned} \tag{3}$$

where $f(t,m,e)$ is a deterministic function of age, education, and marital status, P_t is a permanent income component with innovation N_t, and U_t is a transitory income shock. We assume that lnN_t and lnU_t are normally distributed with education-dependent variances, respectively denoted $\sigma_{N,e}$ and $\sigma_{U,e}$, and means such that their exponent has mean one. Furthermore, lnN_t is allowed to covary with equity returns as detailed below.

At retirement age $t_R = 65$ income becomes non-stochastic and is defined by a replacement rate λ of the age-65 permanent component of income, where λ varies with education and marital status. Thus, $Y_t = \lambda P_{t_R}$ for all $t \geq t_R$.

Assets Agents have two assets in which they can invest: a risk free asset that pays out certain return R_f and a risky equity that pays stochastic return R_t^s. Equity returns are assumed to be lognormal, with mean excess return μ_s. Log equity returns are thus denoted

$$r_t^s - r_f = \mu_s + \epsilon_{s,t}, \tag{4}$$

where $\epsilon_{s,t}$ is distributed normally with standard deviation σ_s, and $corr(lnN_t, \epsilon_{s,t}) = \rho_{n,r}$. The share of savings a household allocates to equities is denoted by α_t. We assume that households cannot hold short positions in either asset, so $\alpha_t \in [0,1]$.

Entry and Participation Costs Households investing in equities pay two types of financial costs. The first time a household invests in equities (i.e., $\alpha_t > 0$), they must pay an entry cost χ. In addition, a per-period participation cost κ is paid in each period an agent allocates non-zero wealth to equity holdings. Participation statuses are denoted as I_t and $Part_t$, where I_t indicates whether a household has ever owned equities and $Part_t$ denotes the current period's participation decision. The total cost of investment is written

$$((1 - I_t) \times \chi + \kappa)) \times Part_t.$$

In our baseline model we assume that costs are constant across the population, but in Section 4.7 we extend the model to allow for entry cost heterogeneity.

Housing We do not formally model housing utility or investment, but follow Gomes and Michaelides (2005) in modeling housing expenditures as an age-dependent mandatory payment expressed as a share of income. Thus, housing expenditures of amount $H_t = h(t)Y_t$ are subtracted from cash on hand at the start of each period.

Lottery Prizes To align the model with our empirical setting, households can receive unanticipated lottery winnings L_t. Households do not form expectations over the prize distribution, meaning that prizes are exogenous and unexpected. Whenever lottery winnings L_t are positive, they enter additively into the budget constraint.

Wealth Accumulation The intertemporal budget constraint is the difference between the sum of income, lottery prizes, and returns on the previous year's non-consumed cash on hand and the sum of housing expenditures and investment costs:

$$\begin{aligned} X_{t+1} = &\left[R_f + \alpha_t(R^s_{t+1} - R_f)\right](X_t - C_t) + \\ &Y_{t+1}(1 - h_t) - ((1 - I_t) \times \chi + \kappa)\, Part_t + L_{t+1}. \end{aligned} \tag{5}$$

Decision Problem and Model Solution For a fixed education and marital status, agents choose consumption C_t, savings $X_t - C_t$, participation status $Part_t$, and portfolio composition α_t subject to age t, prior participation status I_t, wealth X_t, and permanent income P_t. The full specification of the household decision problem is given in Appendix A.1, with model solution details presented in Appendix A.3. As is standard, we exploit the model's homotheticity to eliminate P_t as a model state, and subsequently use lower case variables to reflect normalized states and controls.

Table 5: First-Stage Calibration. This table presents parameter values that are determined separately from our structural estimtion procedure.

Asset Returns

Equity Mean: $\mu_s = .04$ Equity Risk: $\sigma_s = .21$ Risk Free Return: $r_f = .02$

Income Processes by Education Level

	At most High School ($e = 0$)	Some College ($e = 1$)	College or Higher ($e = 2$)
Transitory Risk: σ_U	.188	.188	.205
Permanent Risk: σ_N	.110	.106	.110
Equity correlation: $\rho_{n,r}$	.002	-.001	-.008
Rep. Rate, Single: λ_{t_R}	.685	.641	.617
Rep. Rate, Married: λ_{t_R}	.644	.608	.589

4.2 First Stage Calibration

Table 5 presents parameters calibrated externally from the model. We assume age-specific survival probabilities estimated from observed mortality rates (Appendix A.4 presents details and survival probabilities). Housing expenditures are calibrated to be 30% of income while working, and 20% of income in retirement.

We set the risk free rate to $r_f = .02$, and the excess return and standard deviation on equities to $\mu_s = .04$ and $\sigma_s = .21$, respectively. The assumed equity premium is thus approximately 4.4%, below the historical 6.5% equity premium in Sweden (see Waldenström (2014)). Calibrating a lower than historically observed equity premium reflects unmodeled asset management fees and transaction costs which are estimated to reduce returns to Swedish households by 2% (Calvet et al. (2007)).[11]

We estimate stochastic processes for non-capital income separately by education group as described in Appendix A.5 to obtain average income profiles $f(t, m, e)$ and education-specific income innovation parameters $\sigma_{U,e}$, $\sigma_{N,e}$, and $\sigma_{N,R,e}$. Appendix Figure A.2 shows that average in-

[11]This calibration approach is common in the literature. However, because asset management fees are explicitly accounted for in this calibration, participation costs κ should be thought of as excluding investment fees.

come profiles are hump-shaped and differ in level across education groups. In contrast, Table 5 shows that income innovation parameters are similar across education groups. Notably, the estimated correlation between equity returns and permanent income innovations is negligible. Overall, our estimates of income risk are comparable to values estimated in the United States (e.g., Carroll (1997); Gourinchas and Parker (2002); Cocco, Gomes and Maenhout (2005)), although our estimate of permanent income risk σ_N is relatively small. Retirement replacement rates are approximated as proposed in Laun and Wallenius (2015), with details included in Appendix A.6.

4.3 Estimation Methodology

We next estimate preference parameters and participation costs, namely the time discount factor β, risk aversion ρ, intertemporal elasticity of substitution ψ, entry cost χ, and participation cost κ. Hereafter, we refer to this vector of parameters as $\theta = [\beta, \rho, \psi, \chi, \kappa]$. We estimate θ using the empirical policy function (EPF) approach proposed in Bazdresch, Kahn and Whited (2017).

An EPF is an estimate of the relationship between state variables and policy functions in a structural model. EPFs provide useful benchmarks to evaluate model fit and to identify structural parameters by minimizing the distance between approximations of the model-defined policy functions and their corresponding estimates from the data. Formally, the consumption and participation policy functions from our structural model are written as functions of normalized state variables (t, x, I, l)

$$\begin{aligned} c_i &= c(t_i, x_i, I_i, l_i) \\ Part_i &= Part(t_i, x_i, I_i, l_i). \end{aligned} \tag{6}$$

These policy functions are approximated via a semi-parametric regression using an arbitrary sequence of approximating functions $(h_j(t, x, P, I, l))_{j=1}^{J}$ such that

$$\begin{aligned} c_{i,s} &\approx \sum_{j=1}^{J} b_j^C h_j(t_{i,s}, x_{i,s}, I_{i,s}, l_{i,s}) + \eta_{i,s}^C \\ Part_{i,s} &\approx \sum_{j=1}^{J} b_j^{Part} h_j(t_{i,s}, x_{i,s}, I_{i,t,s}, l_{i,s}) + \eta_{i,s}^{Part} \end{aligned} \tag{7}$$

where $s = 0$ denotes the year of the lottery event. We include linear and quadratic terms for continuous variables (t, x), indicator variables for discrete states (I), and a constant.[12] l_s is omitted from our pre-lottery EPFs as $l_{i,s} = 0$ globally, but included as a linear term in years $s \geq 0$.[13] Details on the exact specification of EPFs for all estimation exercises are included in Appendix A.7.

We rely on registry data from Statistics Sweden to construct the variables in Equation 7. All right-hand side variables are observed directly, as is participation. Although not observed directly, consumption can be constructed from the budget constraint as

$$c_{i,s} = \left[R_f + \alpha_s(R^s_{t+1} - R_f)\right] x_{i,s} + \frac{y_{i,s+1} + l_{i,s} - x_{i,s+1}}{\left[R_f + \alpha_s(R^s_{t+1} - R_f)\right]} \tag{8}$$

and permanent income, which normalizes all continuous variables, can be constructed as described in Appendix A.5.

Having defined our EPFs, we follow Bazdresch et al. (2017) and estimate θ using a particular application of indirect inference (Smith (1993); Gourieroux, Monfort and Renault (1993)). Define $\nu_{i,s}$ as a vector of data observations and let $\nu^k_{i,s}(\theta)$ be the corresponding vector of observations from model-simulation $k = 1, ..., K$ given θ. As discussed in Bazdresch et al. (2017), coefficients b_j from Equations 7 are natural identifying moments. Thus, moment conditions are specified as the vector of differences between model-implied and empiricial coefficients:

$$g(\nu_{i,s}, \theta) = \mathbb{E}\left[b_j(\nu_i) - \frac{1}{K}\sum_{k=1}^{K} b_j(\nu^k_{i,s}(\theta))\right]_{\forall j}. \tag{9}$$

Parameter estimates $\hat{\theta}$ are determined by

$$\hat{\theta} = \arg\min g(\nu_{i,s}, \theta)'\hat{W} g(\nu_{i,s}, \theta), \tag{10}$$

where $\hat{W}$ is the optimal weighting matrix estimated using the procedure described by Erickson

[12]Bazdresch et al. (2017) provide more details on choice of approximating functions h_j, including asympotic convergence of the theoretical and empirical policy functions.

[13]Because income processes differ by education and marital status these are also state variables. We do not include these in our baseline EPF specification to maintain model parsimony and symmetry to estimating Equation 1. Subsequent preference parameter estimates may thus be thought of as the average preferences across education and marital groups. Similarly, we only consider linear effects of lottery wins l_s in our baseline estimation, but allow for nonlinear effects of lottery wins later in this section. Qualitative results are similar if we allow for richer and higher order approximating series in our EPFs, but model fit is worse.

and Whited (2002). Specifically, $\hat{W}$ is the inverse of the clustered covariance matrix $\hat{\Omega}$ of $m(\nu_{i,s})$'s stacked influence functions (denoted $\phi_{m(\nu_{i,s})}$):

$$\hat{\Omega} = \frac{1}{IS} \sum_{i=1}^{I} \left(\sum_{s} \phi_{m(\nu_{i,s})} \right) \left(\sum_{s} \phi_{m(\nu_{i,s})} \right)'. \tag{11}$$

Because our moment vector m consists of coefficients from an OLS regression and Equation 11 does not depend on θ, the influence functions (and thus the optimal weighting matrix) need only be calculated once as the standard OLS influence functions from the empirical estimates of b_j.

Our initial estimation only uses observations prior to the lottery event. To implement this estimation strategy, we sample each household in our post-1999 sample in period $s = -4, ..., -1$ (or earliest observed period if first observation $s_i > -4$) and record all state variables (including observed lottery prizes, where $l_i = 0 \, \forall i$ for $s \neq 0$). We use these observations to estimate Equation 7 to generate empirical moments $b(\nu_i)$. To generate the model counterpart, we solve the model for optimal policy functions and then use these functions, the observed state variables, and random variables' laws of motion to simulate the one-period ahead data set. We then estimate Equation 7 using this simulated data-set to recover the 12 coefficients we target in our baseline estimation. We repeat this simulation $K = 5$ times, construct moment conditions as defined by Equation 9, and calculate the objective function defined in Equation 10. We repeat this procedure until the objective function converges to its minimum value.

In subsequent estimations we simulate household responses to lottery wins. When simulating lottery wins, the procedure is the same except we sample households only in period $s = 0$ and simulate responses assuming sampled prize $l_{i,0}$ enters the budget constraint as detailed in Equation 5. We shuffle lottery prizes within prize group $X_{i,0}$ across simulations so that the simulated distribution of lottery prizes corresponds exactly to the observed distribution. In addition, we add lottery-cell fixed effects to Equation 7 as detailed in Appendix A.7. Finally, to evaluate model fit we use the standard Wald test for overidentification as well as the Wald test for external validity proposed by Bazdresch et al. (2017) (see Appendix A.2 for details).

4.4 Structural Estimation with Pre-lottery Data

Our estimation results based on pre-lottery financial decisions are presented Table 6 Column (1). The resulting model fit for the estimated parameter values (the average $b_j(v_{i,s}^k(\hat{\theta}))$) is presented in Appendix A.7. Panel A shows the estimates and standard errors of preference parameters and participation costs. To facilitate comparison to parameter estimates from other studies, we benchmark our results against estimates from two studies: one that uses a similar sample (Fagereng et al. (2017a), hereafter FGG) and one that uses a similar model specification (Cooper and Zhu (2016), hereafter CZ). FGG estimate a model with CRRA preferences using a representative sample from Norway (where institutions are similar to Sweden), while CZ estimate a model with Epstein-Zin preferences and income heterogeneity by education status using an American sample.

Turning to our estimated preference parameters, our estimate of the time-discount factor $\beta = .869$ is lower than is typically assumed in macro models. However, in models similar to ours a lower discount factor is needed to limit wealth accumulation: FGG and CZ estimate discount factors that vary with model specification from $.75 - .83$ and $.76 - .90$, respectively. Our estimates also suggest – again similar to FGG and CZ – that a bequest motive ($b = 5.191$) is needed to slow asset decumulation during retirement. Finally, our estimates of risk aversion ($\rho = 3.162$) and the elasticity of intertemporal subsitution ($\psi = .645$) are comparable to the baseline estimates in CZ ($\rho = 4.409$, $\psi = .601$). Furthermore, because $1/\rho = .316$ is significantly lower than $\psi = .645$, our estimates reject (as does CZ) the time separable CRRA model where $1/\rho = \psi$.

Entry and participation costs are estimated to be modest relative to total wealth. Per-period costs of stock market participation are economically insignificant at only 10 USD per year. The low costs reflect the persistence in equity market participation: if per-period participation costs were higher, a higher fraction of equity market participants would leave equity markets than what we see in the data. The entry cost, which is identified by the entry decisions of nonparticipant households, is estimated to be 3,217 USD. For comparison, FGG estimate per-period participation costs of 65 to 344 USD, while CZ estimate an entry cost of 684 USD and a transaction cost of 1368 USD. The main reason for our higher entry cost estimate compared to FGG and CZ is a difference in the estimation procedure: our entry cost reflects the average cost for nonparticipants in our sample (presumably participant households in our sample had lower costs of entry) instead of the cost required to generate lifecycle participation rates.

Our model's EPF moments reasonably approximate their empirical counterparts. Given the overidentification test has excellent power to detect even small differences between the model and data generating processes (Bazdresch et al. (2017)), it is unsurprising that the standard overidentification test statistic $\chi^2 = 35.08$ is rejected at all significance levels. Despite this rejection, Appendix A.7 shows the model replicates empirical coefficients with reasonable accuracy. As a further credibility check, in Appendix A.8 we compare the model's predicted wealth and par-

Table 6: Structural Estimation Results and Predictions. Column (1) presents results when the model is estimated using only observations prior to the lottery, Column (2) using only observations after the lottery, Column (3) using observations both before and after the lottery, and Column (4) using observations after the lottery to estimate the entry cost distribution presented in Figure 5 assuming other parameters are fixed at the pre-lottery estimates (Column (1)). Panel A presents the estimated parameters, Panel B presents the model's predictions of the effect of lottery wins on participation probability, Panel C presents tests of external fit for the coefficients indicated in Panel B. In all cases the post-1999 sample is used, and the corresponding coefficients from regressions on consumption are presented in Table A.1.

		Pre-Lottery (1)	**Post-Lottery (2)**	**Pre- and and Post-Lottery (3)**	**Pre- and Post-Lottery: Nonlinear (4)**
A. Parameter Estimates ($\hat{\theta}$)					
Time Discounting - β		.869	.902	.896	.869
		(.019)	(.012)	(.006)	–
Bequest - **b**		5.191	1.32	3.106	5.191
		(1.668)	(.688)	(1.700)	–
Risk Aversion - ρ		3.162	2.360	2.342	3.162
		(.097)	(.091)	(.211)	–
IES - ψ		.645	.595	.669	.645
		(.077)	(.070)	(.063)	–
Entry Cost (1000s USD)- χ		3.217	31.262	12.503	–
		(1.668)	(.688)	(.859)	
Participation Cost (1000s USD) - κ		.010	.040	.036	.010
		(.003)	(.003)	(.006)	–
Overidentifying χ^2 (d.f.).:		35.08 (6)	93.32 (10)	1525.59 (22)	–
B. Comparison of Model Predictions to Lottery Effects					
	Estimate				
i. Baseline Effect					
Full Sample (b^{Part}_{L})	.028	.101	.030	.067	.029
ii. Effect by Prior Participation Status					
Nonparticipants ($b^{Part}_{L_{\|I_{-1}=0}}$)	.104	.313	.113	.209	.104
Participants ($b^{Part}_{L_{\|I_{-1}=1}}$)	.002	.000	.000	.000	.000
iii. Effect by Prize Size (USD), Nonparticipants					
$1.5K < L_i \leq 15K$ ($b^{Part}_{L_{1.5-15K\|I_{-1}=0}}$)	-.012	.013	.013	-.012	.006
$15K < L_i \leq 150K$ ($b^{Part}_{L_{15-150K\|I_{-1}=0}}$)	.078	.172	.017	-.016	.080
$150K < L_i \leq 300K$ ($b^{Part}_{L_{150-300K\|I_{-1}=0}}$)	.156	.644	.026	.462	.158
$300K < L_i$ ($b^{Part}_{L_{300K+\|I_{-1}=0}}$)	.359	.953	.591	.976	.357
C. External validity χ^2 (d.f.)					
Coefficients: B.i,ii,iii		1084.53 (7)	–	–	–
B.i		432.42 (1)	–	–	–
B.ii		546.46 (2)	–	–	–
B.iii		441.49 (4)	127.8 (4)	390.34 (4)	–

ticipation profiles to the empirical age-profiles of wealth and stock market participation. These profiles, commonly targeted in the literature, are not targeted in our EPF estimation strategy but are nevertheless matched reasonably well. Our baseline parameter estimates slightly overpredict wealth accumulation early in life and decumulation later in life, but otherwise provide a decent approximation to lifecycle saving and participation patterns.

In Table 6, Panel B we compare the model's predictions of the effect of lottery wins on participation to their empirical counterparts (displayed at the left-hand side of panel B).[14] Panel B.i shows that our model predicts each 150K USD (1M SEK) increases stock market participation by .101 in the full sample, 3.6 times larger than the empirical estimate of .028.[15] Panel B.ii shows that this overprediction is driven by a .313 increase in participation among pre-lottery nonparticipants, which is again much larger than its empirical counterpart. Thus, our baseline estimation predicts responses to lottery wins that are qualitatively consistent with the main results in Section 3, but much larger quantitatively.

Our model's predictions regarding the effect of lottery winnings on stock market participation are intuitive. First, because participation costs are negligible most participant households are predicted to participate regardless of lottery winnings, resulting in a near-zero coefficient on $l_{i,s}$. Second, entry costs are not large enough to discourage nonparticipants who win large prizes from entering the stock market. This can be seen in Panel B.iii: relative to prizes $<$10K USD (the omitted category), prizes from 150-300K USD are predicted to increase participation among pre-win nonparticipants by .644 while the corresponding number for prizes sized 300K+ USD is .953. Both effects far exceed their corresponding empirical estimates of .156 and .359, respectively. In Table 6, Panel C we present the external validity test proposed by Bazdresch et al. (2017) for these three specifications jointly and separately. Unsurprisingly given the poor fit, tests of external validity strongly reject our baseline estimation's match of all lottery predictions.

4.5 Structural Estimation with Lottery Data

To understand what model parameters – in particular entry costs – are needed to account for our lottery results, we re-estimate our model using only participation decisions after the lottery event.

[14] Appendix A.7 details the exact regressions we estimate to obtain the model-predicted lottery coefficients. Also, note that our EPF coefficients on lottery wins differ trivially from the coefficients presented in Section 3 due to differences in specifications defined in Equations 1 and 17.

[15] The empirical moments used in our structural estimation are slightly different than the corresponding estimates presented in Section 3 due to differences in specification between Equation 1 and our EPFs (see Appendix A.7). These differences are generally negligible.

In practice, we estimate the model targeting 16 benchmarks: all coefficients except for cell fixed effects from the post-lottery participation and consumption EPFs, and the lottery coefficients from participation and consumption regressions by pre-lottery participation status. The exact regressions, coefficients, and resulting model fit are presented in Appendix A.7 and Table A.1. The optimal weighting matrix is again calculated as the inverse of the influence function from these regressions.

Column (2) of Table 6 presents the results from this estimation. Preference parameters overall are estimated to be similar to our estimates from pre-lottery data: agents are estimated to significantly discount the future ($\beta = .902$), have a reasonably strong bequest motive ($b = 1.32$), moderate risk aversion ($\rho = 2.36$) and high IES ($\psi = .595$). Furthermore, per-period participation costs are again estimated to be negligible at only 40 USD per year. Thus, most estimates of parameters from lottery data are consistent with those from pre-lottery observations.

Entry costs are however estimated to be 31,262 USD, an order of magnitude larger than our baseline estimate. This cost is quite significant economically and corresponds to approximately 30% of average wealth or 70% of annual income in our sample. It is difficult to reconcile such a high cost of entry with any reasonable cost that households might pay to enter equity markets. However, the intuition behind these large costs parallels the intuition behind the lottery predictions of our baseline estimates: matching low rates of equity market entry after receiving large lottery prizes requires a large disincentive. In the model this disincentive is entry cost χ, which needs to be extremely large to match our lottery estimates.

The standard overidentification test statistic $\chi^2 = 98.23$ is rejected at all significance levels. A main reason for this rejection is that the model is unable to generate MPCs from lottery wealth as high as those observed in the data and still match the consumption policies of households that did not win large lottery prizes. Despite the statistical rejection, Appendix A.7 shows that the model generally matches the empirical coefficients. Furthermore, Appendix A.8 shows that the model reasonably approximates lifecycle wealth profiles, although large entry costs reduce entry over the lifecycle to virtually zero. Finally, we test and reject the external validity of the model's predicted nonlinear effects of lottery win on participation in Panel C. A one-time cost of 31K USD disincentivizes virtually all winners except those of more than 300K USD (2M+ SEK) from entering equity markets, while empirical estimates suggest larger effects on winners of smaller prizes and smaller effects on winners of larger prizes. Thus, the model has difficulty replicating effects on consumption and prize size heterogeneity with a single large cost despite being able to replicate other effects of lottery wins on stock market participation.

4.6 Structural Estimation with Pre-Lottery and Lottery Data

An appropriate model of stock market participation should be able to account for participation decisions both before and after lottery wins. Therefore, we next estimate our model targeting the combined pre- and post-lottery benchmarks matched separately in Columns (1) and (2) of Table 6. In practice, we stack the two moment vectors from our previous two estimations, and re-estimate Equation 10 with the optimal weighting matrix defined by the inverse of the covariance matrix of these stacked influence functions. The resulting estimates are presented in Column (3) of Table 6.

Panel A shows the parameter estimates are, not surprisingly, similar to those from targeting pre-lottery and lottery coefficients separately (Columns (1) and (2), respectively). The parameter of main interest, the one-time entry cost, is estimated to be 12,503 USD. This estimate is thus closer to the the baseline estimate of 3,217 USD than the lottery estimate of 31,262 USD. Intuitively, the optimal weighting matrix assigns more weight to better identified moments and given that the standard error on the prior participation coefficient in the baseline participation EPF is lower than the standard error on lottery winnings in our lottery EPF, thus assigns greater weight to pre-lottery coefficients. However, including the lottery estimates and their higher implied barriers to entry does increase the entry cost estimates by over 9K USD relative to baseline.

The overidentification test shown in Panel A indicates that the model's fit is strongly rejected. This is due to poor fit of both pre-lottery and lottery moments (see Appendix A.7). The model-predicted effects of lottery wins on participation presented in Column (3) generally fall between the pre-lottery and lottery predictions in Columns (1) and (2), and thus overpredict the effects of lottery prizes on participation. Furthermore, a test statistic of external validity of the nonlinear fit presented in Panel C are rejected at all significance levels. Thus, a model where participation is determined by costs of entry and participation is a poor approximation of household financial decisions.

4.7 Structural Estimation with Entry Cost Heterogeneity

The model's predictions regarding effect by prize size are soundly rejected (Panel B.iii) in all of the considered structural estimations, including when we only target lottery wins (Column (2)). A fundamental economic question is thus still unanswered: What size and structure of participation disincentives are required for our model to match the full distribution of household participation responses to lottery wins?

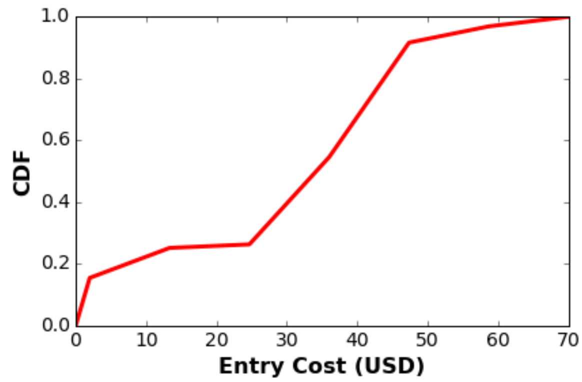


Figure 5: Structural Estimates of Entry Costs.

In Column (4) we conduct a final structural exercise. We hold preference parameters and per-period participation costs fixed at their baseline pre-lottery estimates, but extend the model to allow for heterogeneity in entry costs as determined by the cost distribution

$$\chi_i \sim G_\chi(x). \tag{12}$$

We approximate this distribution by seven equi-distant points between 2K and 70K USD and estimate the the probability mass for each point of this distribution. Our moment vector includes the estimated effects of lottery winnings on participation, namely the seven coefficients from Panel B.i-iii of Table 6.[16] Thus, the resulting entry cost distribution reflects the entry disincentives needed to match our lottery estimates (in particular the effects by prize size) given preference parameters implied by pre-lottery estimates.

In Column (4), Panel B we observe that model predictions match their empirical counterparts nearly exactly. This is not surprising given the flexibility of the assumed entry cost distribution. The resulting estimated cost CDF G_χ is presented in Figure 5. These estimates imply that approximately 23% of our sample have entry costs $\leq$ 10K USD while approximately 40% have entry costs $\geq$ 40K USD, suggesting that our previous structural estimates masked significant heterogeneity in entry costs. In addition, the mean and median implied costs of entry are estimated to be quite large, at 43K USD and 36K USD respectively, confirming that large disincentives are needed to account

[16]The simulation procedure is exactly the same as above, except for each simulation we sample an entry cost for each household from this distribution. Furthermore our estimation procedure is unchanged – we estimate Equation 10 with optimal weighting matrix determined by the stacked influence functions - although test statistics are invalid given this system is just identified.

for the observed effects of lottery wealth on participation probability. Furthermore, the shape of the estimated entry cost distribution mirrors the empirical estimates of effects by prize size. Accounting for positive effects of lottery wins on entry for small and intermediate prize winners requires some households to have small entry costs, while matching the small rates of entry of winners in our largest prize categories requires a majority of housholds to have large entry costs.

5 Explaining Nonparticipation

The previous section showed that accounting for the causal estimates analyzed in Section 3 requires either implausibly large entry costs or a richer model of stock market participation than fixed costs of entry and participation. To better understand what types of disincentives might explain the puzzingly small effects of wealth on stock market participation, we present a series of additional reduced-form and structural analyses that examine various pathways' credibility in explaining our results. We examine three broad classes of pathways: alternative economic explanations, alternative preference specifications, and pathways related to expectations and information processing. We focus on pre-lottery nonparticipants in this section, while corresponding estimates for pre-lottery participants are presented in Appendix Tables B.8, B.9, B.10, and B.11.

Our reduced-form analyses fall into two categories. We examine how lottery winnings affect investment in other assets besides equities. We also examine treatment effect heterogeneity by comparing effects across different population strata. While heterogeneity analyses are informative, they are subject to the caveat that only wealth is randomly assigned. Thus, we cannot be certain that heterogeneity in treatment effects is directly caused by the dimension being examined and not another factor that is correlated with that dimension.

5.1 Alternative Economic Explanations

Net wealth Pre-lottery nonparticipants with higher wealth have more financial resources and thus stronger incentive to participate in equity markets. Given that stock market participation is strongly predicted by wealth in the cross section, wealthy households not already participating in equity markets might face different entry disincentives than poorer nonparticipants. In particular, nonparticipation by lottery players with low pre-lottery wealth could plausibly be explained by costs of entry even smaller than those estimated from pre-lottery data (Table 6, Column (1)).[17] Hence,

[17] Our structural analysis controls for the pre-lottery wealth distribution by sampling households at their observed states but focuses on the treatment effect of the average pre-lottery nonparticipant. Presumably, allowing entry costs to vary with wealth would result in higher entry costs for nonparticipants with more pre-lottery wealth.

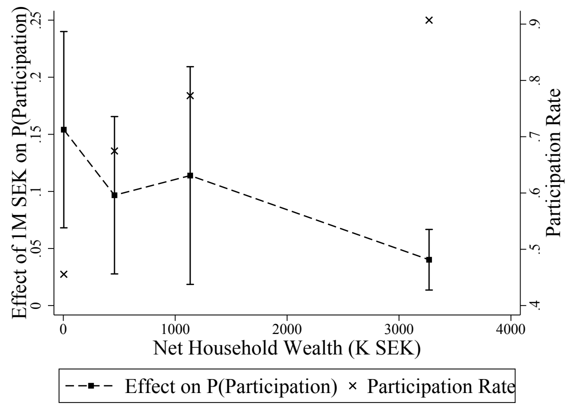
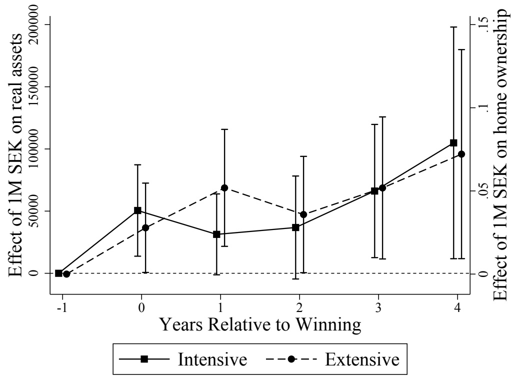


Figure 6: Effect of Wealth (1M SEK) on Participation by Pre-lottery Wealth. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 separately for each pre-lottery wealth quartile. See Appendix Tables B.11 for the underlying estimates and results for $s = -1$ equity market participants.

if the effect of lottery wealth on entry is much larger for households with low pre-lottery wealth, this might suggest that a model with modest fixed costs is a good approximation for low-wealth households and that the difference between model predictions and estimated effects is driven by high-wealth households that had already revealed themselves unwilling to enter equity markets.

Figure 6 plots the effect of each 150K USD (1M SEK) on stock market participation probability by pre-lottery wealth quartile on the left axis, as well as the pre-lottery participation probability on the right axis. Stock market participation is strongly increasing in wealth, with a pre-lottery participation rate among the richest quartile of households (92.1%) approximately twice that of households in the bottom quartile (46.9%). Below median wealth households thus comprise most of our nonparticipant sample (76.2%) and contribute more to the average effect among nonparticipant households (Figure 2, Panel (a)). We also find that the effect of lottery wins on participation decreases with pre-lottery wealth, particularly among the wealthiest 25% of households. In the top wealth quartile, each 150K USD (1M SEK) increases entry probability by only .040, suggesting that high wealth nonparticipants are not responsive to large windfall gains.

There are two key takeaways from these results. First, wealthy nonparticipants that had previously revealed themselves unwilling to enter equity markets are largely unaffected by windfall gains. This confirms the previously documented “significant challenge to financial theory” posed by nonparticipation of wealthy households (Campbell (2006), p.1564): accounting for their behavior requires moving beyond models of financial costs and gains. Second, because the estimated



(a) Real Estate

(b) Debt

Figure 7: Effect of Wealth (1M SEK) on Real Estate and Debt Investment. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 in the post-1999 sample of $s = -1$ equity market nonparticipants. Panel (a) shows the effect of 1M SEK (150K USD) on real estate market participation probability and the effect of 1 SEK on net real estate wealth. Panel (b) shows the effect of 1M SEK (150K USD) on probability of becoming debt-free and the effect of 1 SEK on total debt. See Appendix Tables B.9 and B.10 for the underlying estimates and results for $s = -1$ equity market participants.

effect is much smaller than model predictions for households in lower wealth quartiles as well, this challenge extends to most non-wealthy households too. Thus, there are few households for whom nonparticipation can be accounted for without alternative explanations.

Investment in Illiquid Assets Illiquid, high-returning assets like real estate might have two distinct effects on investment decisions. First, they provide alternative investment opportunities which may crowd-out equity purchases (Grossman and Laroque (1990); Mayer and Engelhardt (1996); Cocco (2005); Flavin and Yamashita (2011); Vestman (2013); Fuster and Zafar (2014)). Second, liquidity-constrained households might experience lower welfare gains from stock market participation (Kaplan and Violante (2014); Campanale, Fugazza and Gomes (2015)).

To explore whether real estate purchase crowds out stock market entry, Figure 7 presents the estimated effect of wealth on extensive and intensive real estate investment among pre-lottery equity market nonparticipants. We find that each 1M SEK increases the probability of owning real estate by .028 the year of and .052 one year after the lottery event.[18] We also find small effects on intensive margin investment in real estate – real estate wealth increases by 4.5% of the amount

[18]In unshown analyses we find that these effects are larger – .074 the year of and .139 one year after the lottery event – when further conditioning on not owning real-estate prior to winning.

won the year of and 2.8% one year after the lottery.

To explore whether liquidity constraints affect the decision to enter equity markets, in Table 7 we present treatment effects conditional on liquid wealth share (as measured the year prior to the lottery event). We find no difference among households with above and below median liquid asset share, suggesting that liquidity is not a major determinant of stock market entry among pre-lottery nonparticipants. Similarly, Table 7 shows that there is no significant difference in effects on equity market entry among households that do and do not own real estate. Taken altogether, there is limited evidence that the small effect of wealth on stock market entry is explained by investment in illiquid assets or real estate.

Debt Besides real estate investment, paying down debts in lieu of equity market entry may be optimal if interest rates on debt are substantially higher than the risk-free rate (Davis et al. (2006); Becker and Shabani (2010)). Table 7 therefore compares estimated effects on equity market entry conditional on having and not having outstanding debt. The estimated effect on participation is about twice as large in debt-free households, suggesting paydown of debt as a credible explanation for indebted households that did not participate in equity markets prior to the lottery event.

To further test this explanation, Panel (b) of Figure 7 estimates the effect of lottery wealth on outstanding debt and the probability of becoming debt free among households that did not participate in equity markets and had positive debt prior to the lottery event. We find that lottery wealth has a small effect on outstanding debt, with average debt reduction being insignificant and never exceeding 5% of the amount won at all horizons. We similarly find modest effects on the probability that households exit debt markets, with the effect of 1M SEK on debt-market exit peaking at 10% of the amount won the year after the lottery event.

Thus, despite the heterogeneity in treatment effects among households that do and do not have debt, we find no clear evidence that investment in debt reduction is a significant use of lottery winnings. The differential effects therefore likely reflect a common correlation with other factors that shift participation incentives, not the effect of debt itself.[19]

Income Risk/Self Employment The structural model in Section 4 accounts for historically observed income risk processes, but average income profiles (conditional on education) might not adequately capture risks for all households. For example, self-employed workers likely benefit less from stock market participation due to both higher uninsurable wage risk and higher corre-

[19] In unshown analyses we examine predictors of having no debt. We find that age, education, and income, are significant predictors of having no debt.

Table 7: Heterogeneous Effect of Wealth (1M SEK) on Participation Probability among $s = -1$ Equity Market Nonparticipants, Alternative Economic Channels. Coefficients are obtained by estimating Equation 1 at time $s = 0$ in the post-1999 sample of equity market nonparticipants at time $s = -1$. Hetero p obtained from an F-test of the null hypothesis that the two lottery-wealth coefficients are identical. See Appendix Table B.8 for results for time $s = -1$ equity market participants.

	Liquid Wealth		**Home Owner**		**Gross Debt**	
	Low	**High**	**No**	**Yes**	**=0**	**>0**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Effect	.120	.129	.147	.105	.212	.092
SE	.028	.041	.052	.027	.037	.025
p	.000	.002	.005	.000	.000	.000
Hetero p	.855		.474		.007	
N	9,935	9,343	8,022	11,256	9,545	9,733
$\% \, Part_{s-1}$	.717	.726	.554	.784	.679	.759
	Self-Employed		**Net Wealth**		**Age**	
	No	**Yes**	**Low**	**High**	**$\leq$ 61**	**(> 61)**
	(7)	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Effect	.131	.046	.137	.035	.118	.128
SE	.026	.040	.028	.026	.027	.046
p	.000	.246	.000	.185	.000	.006
Hetero p	.079		.008		.838	
N	18,628	650	15,706	3,572	8,495	10,783
$\% \, Part_{s-1}$	.719	.832	.601	.884	.759	.690

lations between income shocks and equity returns (Heaton and Lucas (2000b); Viceira (2001); Fagereng, Guiso and Pistaferri (2017b)).

Table 7 compares the treatment effect for households that did and did not have self-employment income prior to the lottery event. The estimated effect for self-employed is small and statistically insignificant, suggesting non-entry by self-employed nonparticipants can be explained through this channel. However, less than 3.4% of all nonparticipants had self-employment income, implying that self-employment only can explain continued nonparticipation for a small fraction of pre-lottery nonparticipants.

Age The accumulated gains from investing in equities not only increase in initial investment, but also in the investment horizon for which households foresee accumulating returns. Thus, if entry costs or upfront disincentives contribute to the participation decision one would expect younger households to be more likely to enter equity markets than older households. To test for this behavior, in Table 7 we examine treatment effects among pre-lottery nonparticipants above and below median age (61) and find no evidence that younger households are more likely to enter.

Restricted Subsample Our results thus far indicate that several unmodeled participation disincentives are unlikely to account for the model's overprediction of the effect of lottery prizes on participation. To further check that the combined effects of several of these disincentives do not explain our results, we next compare reduced-form estimates and model predictions for a subsample of households for whom the model might more accurately reflect key determinants of stock ownership. We restrict our sample to households with no self-employment income, debt below 15K USD (100K SEK), and net wealth below 1M USD (6.7M SEK). Table 8, Columns (d) and (e) presents the empirical estimates and model predictions for this subsample. The model predictions are generated using the parameter estimates obtained using pre-lottery data (Table 6, Panel A, Column (a)).

In this restricted sample, the model predicts that each 150K USD (1M SEK) increases participation by .107, while empirical estimates show an increase of only .039. This difference is again driven by pre-lottery nonparticipants: the model predicts an increase of .286 compared to the estimated effect of only .152. These differences between model predictions and empirical estimates are smaller than the differences for the full sample (Table 6, Column (a), Panel B), but the improvement in model fit is relatively small. Thus, even when comparing to a sample whose financial decisions more closely align with those that are modeled, a significant gap between model predictions and empirical estimates remains.

5.2 Preference-based Explanations

Apart self-employment – which affects only a small percentage of our nonparticipant sample – we find limited evidence that the examined alternative economic pathways account for the low rates of equity market entry. We next examine the credibility of channels related to alternative preference specifications.

Status Quo Bias We conduct two analyses to test whether status quo bias, or a general reluctance to invest in new financial products, can explain low equity market entry. Because lottery winnings are paid directly to a bank account, investing in other asset classes requires an active decision. An indication of status quo bias is therefore if a large share of the amount won remaining in the bank

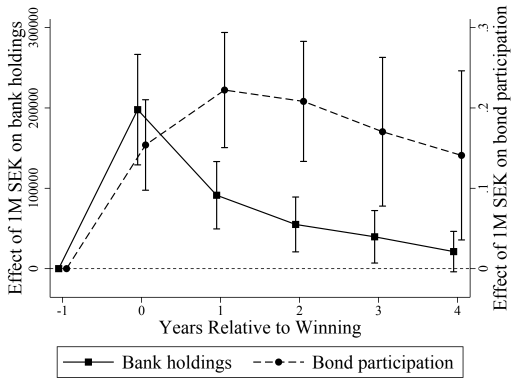
Table 8: Structural Model Predictions, Alternative Specifications and Calibrations. Columns (1)-(3) present results from the post-1999 sample including estimates, model predictions after the model is re-estimated assuming a present bias parameter $\beta = .6$, and model predictions assuming the subjective equity premium matches the surveyed distribution presented in Figure 9 with other parameters given by pre-lottery estimates (Table 6, Column (1)). Columns (4) and (5), respectively, present estimates and model predictions assuming parameters from pre-lottery estimates after restricting the sample to households with no self-employment income, debt less than 15K USD, and net wealth less than 1M USD. Columns (6) and (7) present estimates and model predictions assuming parameters from our pre-lottery estimates if the sample is restricted to winners with some secondary education and above median IQ.

	Full Sample			Restricted Finances Subsample		High Information Subsample	
	Est. (1)	Present-Bias (2)	Subj. Returns (3)	Est. (4)	Model (5)	Est. (6)	Model (7)
i. Baseline Effect							
Full Sample (b_L^{Part})	.028	.115	.066	.039	.107	.013	.067
ii. Effect by Prior Participation Status							
Nonparticipants ($b_{L\|I_{-1}=0}^{Part}$)	.104	.340	.197	.152	.286	.163	.248
Participants ($b_{L\|I_{-1}=1}^{Part}$)	.002	.000	.000	-.006	.000	-.026	.000
iii. Effect by Prize Size, Nonparticipants							
$10K < L_i \leq 100K$ ($b_{L_{10-100K\|I_{-1}=0}}^{Part}$)	.-.012	-.012	-.014	-.043	.032	–	–
$100K < L_i \leq 1M$ ($b_{L_{100K-1M\|I_{-1}=0}}^{Part}$)	.078	.097	.003	.011	.042	–	–
$1M < L_i \leq 2M$ ($b_{L_{1-2M\|I_{-1}=0}}^{Part}$)	.156	.680	.452	.247	.471	–	–
$2M < L_i$ ($b_{L_{2M+\|I_{-1}=0}}^{Part}$)	.359	.963	.709	.943	.895	–	–
		N=70,139		N=44,561		N= 3,355	

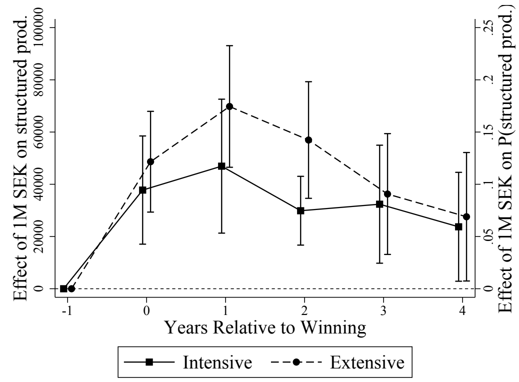
account. In Figure 8, Panel (a) (left axis) we show that bank account balances increase by 20% of the amount won in the year of the lottery event (for reference, total wealth increases on average by 70% the amount won), and then falls quickly with time since the lottery. We interpret the limited effect on the bank account balance as evidence against status quo bias.

Another potential manifestation of status quo bias is a bias towards investing winnings in asset classes in which households had previously invested. To test this, Figure 8, Panel (a) (right axis) shows the estimated effect of lottery wealth on bond ownership. The effect on owning bonds is approximately twice as large as the effect on owning equities at all horizons.[20] Thus, inconsistent

[20]Results regarding the effect of lottery wealth on bond ownership are similar when we restrict to the sample of



(a) Bank/Bond



(b) Structured Products

Figure 8: Effect of Wealth (1M SEK) on Bank/Bond and Stuctured Product Investment. Coefficients and 95% confidence intervals are obtained by estimating Equation 1 in the post-1999 sample of $s = -1$ equity market nonparticipants. Panel (a) shows the effect of 1M SEK (150K USD) on bond market participation probability and the effect of 1 SEK on total bank account balances. Panel (b) shows the effect of 1M SEK (150K USD) on probability of owning structured products and the effect of 1 SEK on total structured product holdings. See Appendix Tables B.9 and B.10 for the underlying estimates and results for $s = -1$ equity market participants.

with status quo bias, lottery winnings induce households that do not participate in equity markets to invest in some new financial products.

Loss Aversion and Structured Products Loss aversion, a preference specification in which individuals are more sensitive to losses than gains around a reference point (Tversky and Kahneman (1986)), is a commonly proposed explanation for limited equity demand (e.g., Berkelaar, Kouwenberg and Post (2004); Ang, Bekaert and Liu (2005); Barberis, Huang and Thaler (2006)) supported by empirical evidence (Dimmock and Kouwenberg (2010)).

During the period of our study, retail structured products became increasingly important assets in Sweden, with ownership increasing from approximately 0% of households in 2000 to 17% by 2007. These assets were particularly popular among households that traditionally did not participate in equity markets. Nearly all structured products (98%) offered capital protection against downside risk (Calvet, Célérier, Sodini and Vallée (2017)). Given the suitability of structured products for loss-averse households, one might expect to see large increases in holdings of these assets following lottery wins if loss aversion explains small effects on equity market entry.

households that neither owned equities nor bonds prior to the lottery event.

Figure 8 presents the effect of lottery wealth of structured product investment. On the extensive margin, we find that 150K USD (1M SEK) increases structured product ownership by 10-17 percentage points, depending on the time since the lottery.[21] However, intensive-margin investment in structured products is small and never exceeds 5% of the total amount won. Furthermore, in unshown analyses we found that roughly one half of nonparticipants who entered the structured product market also entered equity markets. Thus, some nonparticipating households purchase assets with downside protection, but this does not explain the small effect of wealth on stock market entry for most pre-lottery nonparticipants.

Present-Biased Preferences Our empirical and structural results suggest up-front disincentives to equity market entry are large, while continued participation disincentives are small. An alternative preference specification consistent with this disincentive structure is present-biased time preferences which might make households less willing to pay entry costs and forego consumption today in exchange for higher future consumption.

To test whether present bias can account for our results, we extend our model to allow for naive quasi-hyperbolic discounting in the form of $\beta - \delta$ time preferences (Laibson (1997)). We re-estimate the structural model using pre-lottery data and assuming a present-bias parameter of $\beta = .6$.[22] Table 8, Column (2) shows that even after accounting for present-bias in this manner, the model still significantly overpredicts the effect of lottery wins on participation by an amount comparable to our pre-lottery estimates.[23]

5.3 Information and Belief-based Explanations

A third broad type of explanation for continued nonparticipation is that households have information or beliefs that lead them to undervalue participation in equity markets.

Education Cross-sectional analysis generally shows a positive correlation between education and stock market participation, consistent with more educated households having better under-

[21] Appendix Table B.7 shows that each 150K USD (1M SEK) increases participation by 18 percentage points among nonparticipants the year of the win when participation is defined as owning either equities or structured products.

[22] The assumed value of present-bias is consistent with experimental evidence in Angeletos, Laibson, Repetto, Tobacman and Weinberg (2001) and is used by Love and Phelan (2015) in exploring the role of quasi-hyperbolic discounting in a lifecycle model with Epstein-Zin preferences.

[23] Comparing Table 8, Panel iii with Table 6, Panel B.iii, the present-bias model better explains the behavior of small and medium prize winners, but fails to account for the for non-entry among large prize winners. Given that our experiment focuses on large wealth shocks, modeling impatience via present-biased preferences only marginally reduces the difference between model-predicted and empirical effects.

Table 9: Heterogeneous Effect of Wealth (1M SEK) on Participation Probability among $s = -1$ Equity Market Nonparticipants, Belief and Information Channels. Coefficients are obtained by estimating Equation 1 at time $s = 0$ in the post-1999 sample of equity market nonparticipants at time $s = -1$. Hetero p obtained from an F-test of the null hypothesis that the two lottery-wealth coefficients are identical. Equity returns are based on the MSCI Sweden Index the calendar year prior to the lottery. See Appendix Table B.8 for results for time $s = -1$ equity market participants.

	Education		**Cognitive Skill**		**Recent Equity Returns**		**Early Equity Returns**	
	No College (1)	**College (2)**	**Low (3)**	**High (4)**	**Low (5)**	**High (6)**	**Low (7)**	**High (8)**
Effect	.107	.223	.039	.304	.053	.140	.086	.176
SE	.025	.053	.055	.147	.039	.028	.030	.036
p	.000	.000	.476	.038	.167	.000	.004	.000
Hetero p	.050		.090		.069		.056	
N	16,510	2,768	804	957	10,402	8,876	10,591	8,687
$\% \, Part_{s-1}$	.686	.842	.677	.790	.742	.703	.721	.730

standing of equities, facing smaller information costs, or experiencing less psychological discomfort from owning stocks (Van Rooij, Lusardi and Alessie (2012); Benjamin, Brown and Shapiro (2013)). Table 9 shows that the treatment effect indeed increases with education: each 150K USD (1M SEK) causes an increase in entry probability of .223 among households with a college degree, approximately twice as large as the .107 increase among households without a college degree. Thus, higher educated households are more likely to enter equity markets due to wealth increases, suggesting that high entry cost estimates potentially reflect information or psychological costs.

Cognitive Skills Similar to education, previous research has documented a positive correlation between stock market participation and cognitive skill (e.g., Grinblatt, Keloharju and Linnainmaa (2011)). To test for heterogeneity by cognitive skill, we consider a subsample of lottery players for whom physical and psychological traits were measured for mandatory enlistment (for all young men) in the Swedish military.[24]

Table 9 presents treatment effects among pre-lottery nonparticipants stratified by above and below median cognitive ability. Winners with above median cognitive ability are both more likely to participate in equity markets before the lottery event and to enter the stock market conditional

[24]See Lindqvist and Vestman (2011) for a detailed description of the cognitive skill measure from the Swedish military draft.

on pre-lottery nonparticipation. Each 150K USD (1M SEK) increases entry probability for males of high cognitive ability by .304 as opposed to .047 for males of low cognitive ability. This large difference supports the importance of pathways related to information processing in accounting for the low rates of stock market entry following a large windfall gain.

Recent Equity Returns Next we examine whether the effect differs among households that experience lottery wins following years of positive equity returns, as it may if individuals overweight recent events when forming beliefs about future returns. Previous survey work has found investors are overly pessimistic (optimistic) following periods of negative (positive) market returns (Vissing-Jørgensen (2003); Greenwood and Shleifer (2014)), suggesting that individuals do extrapolate recent returns when forming beliefs. Consistent with this hypothesis, Table 9 shows that each 150K USD (1M SEK) increases participation probability by .140 for households that win in years following positive equity returns, while the effect is only .056 for households that win following negative equity returns.[25]

Early Equity Returns In addition to recent equity return experiences, Malmendier and Nagel (2011) show that equity returns during formative years positively correlate with stock market participation much later in life. Furthermore, they note that negative early equity return experiences predict more pessimistic beliefs. More recently, Fagereng et al. (2017a) show that equity returns experienced during ages 18-25 are a good proxy for cohort effects when estimating stock market participation profiles.

To check whether early equity return experiences affect willingness to enter the stock-market following a windfall gain, we stratify our sample according to average equity returns experienced between ages age 18 and 25. Consistent with the importance of formative years, Table 9 shows that each 1M SEK increased stock market entry probability by .176 among nonparticipants who experienced above median returns compared to .086 among those who experienced below median returns.

Subjective Beliefs Regarding Equity Returns We next consider whether subjective beliefs about equity returns can explain the small effects on equity market entry. This analysis relies on responses to a survey fielded to 4,820 members of our lottery sample in Fall 2016. The survey asked lottery winners about (1) the probability that the Stockholm Stock Exchange index would increase in value during the next 12 months and (2) the expected increase or decrease in the Stock-

[25] In unshown analyses we confirm that very similar patterns hold at quarterly frequency.

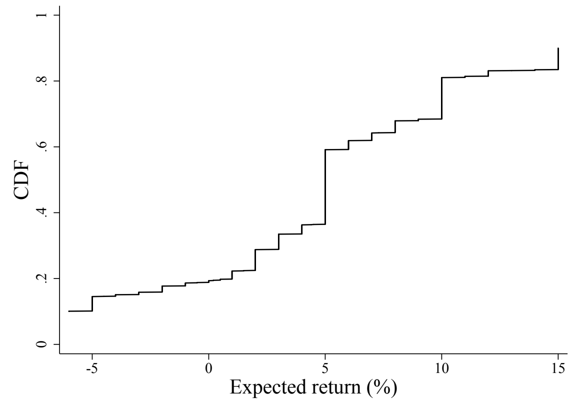


Figure 9: Subjective Distribution of Equity Returns. The above figure presents the population CDF of survey respondents' expected market returns during the 12 months following the survey (i.e., Fall 2016-Fall 2017). For expositional purposes, we truncate the distribution at the 10^{th} and 90^{th} percentiles.

holm Stock Exchange index during the next 12 months. We refer to Appendix C for a detailed account of the survey procedure and a complete analysis of the survey responses.

Our analysis of survey responses yields several insights. First, as shown in Figure 9, there is substantial heterogeneity in expected stock returns with a large fraction of respondents reporting pessimistic beliefs. Although the average expected return of 5.9% is only slightly below the historical average of 8.5%, 67.9% of respondents report beliefs below the historical average, and 18.8% of respondents report a negative expected return.[26] Second, approximately 40% of respondents report responses to our two survey measures that appear mutually inconsistent, suggesting that numeracy and financial literacy is low in our sample. Thus, for a substantial proportion of our sample pessimistic beliefs and lack of financial literacy could contribute to low rates of stock market entry (Van Rooij et al. (2012)).

Given the pessimistic beliefs about stock market returns, our previous structural estimation exercises – where beliefs are calibrated to historical realizations – likely overstates the perceived participation incentives for a significant proportion of households. To test whether pessimistic beliefs can explain our results, we simulate model predictions using the subjective equity return belief distributions in Figure 9. We assume parameters from our pre-lottery estimates from Section

[26]Such findings are not unique to our study. Hurd (2009) and Dominitz and Manski (2011) document substantial heterogeneity in equity return beliefs, with many housholds holding equity return beliefs substantially more pessimistic than historical data would suggest. In fact, Hurd (2009) concludes that equity returns are sufficiently pessimistic for enough households to account for observed stock market nonparticipation.

4.4, randomly sample an assumed equity premium from this distribution, and then simulate model predictions of causal effects of wealth on stock market participation. The resulting predictions are presented in Column (3) of Table 8.

Under this calibration, the model predicts that each 1M SEK causes an increase in participation probability of .197 for pre-lottery nonparticipants, resulting in a predicted overall effect of .066. Accounting for subjective beliefs thus make progress in aligning the model predictions with our empirical estimates, reducing the difference by about 50%.

High Information Subsample Our research design is not well suited to rigorously discriminate between different models of information aquisition and belief formation in household portfolio choice (e.g., Van Nieuwerburgh and Veldkamp (2010); McKay (2013); Lusardi, Michaud and Mitchell (2017)). However, if differences between model-assumed and actual expectations contribute to our results, it is reasonable to expect empirical and model estimates to be more aligned for households with expectations more closely aligned with those specified in the model.

To test this hypothesis, we examine model predictions for a sample of households that are likely better able to acquire and process information. We restrict the sample to households with at least some secondary education and where the winner had above median IQ. Using pre-lottery parameter estimates, in Table 8 we compare model-implied estimates in Column (6) to actual estimates in Column (7) for this sample.[27]

Table 8 shows that for our high information sample model-predicted and estimated effects align reasonably well. In particular, the model predicts an increase in participation probability of .248 per 1M SEK, which aligns reasonably well the empirical estimate of .163. Restricting our sample to high-information households thus reduces the difference between model-predicted and empirical effects from .209 to .085 in our full sample (Table 6, Column (1)).

6 Conclusion

The widespread nonparticipation in the stock market observed across Western countries is a much studied but imperfectly understood phenomenon (Vissing-Jørgensen (2003); Campbell (2006); Guiso and Sodini (2013)), partly because stringently testing theoretical predictions is often challenging in observational data. In this paper, we leverage the randomized assignment of wealth in three Swedish lotteries to estimate the causal effect of wealth on stock market participation.

[27] We do not consider effects by prize size in this estimation because the sample size is to small to generate credible predictions.

In our main reduced-form analyses, we find that each 150K USD (1M SEK) causes an immediate increase in stock market participation by 12 percentage points among pre-lottery nonparticipants, but has no effect on pre-lottery participants. Furthermore, there is no effect on participation if lottery wins are paid out monthly instead of as a lump-sum, and even among households that receive more than 300K USD a majority of pre-lottery nonparticipants do not enter equity markets.

While clean causal estimates are important, it is difficult to interpret these estimates without a framework based in economic theory. To address this concern, we introduce and estimate a rich structural life-cycle model that replicates the financial decisions which households make before and after a windfall gain. Using structural parameter estimates from targeting only non-experimental data, we find that the model significantly overpredicts the effect of a large windfall gain on stock market participation: disincentives to enter equity markets are not large enough to keep a majority of large prize winners out of equity markets. Furthermore, matching the causal estimates requires entry cost estimates that are far too large to be reasonably interpreted as financial costs. This suggests that that the "nonparticipation puzzle" extends well beyond the wealthiest stock market nonparticipants, is larger than has been previously documented, and that any theory of stock market participation consistent with our results will include a large, up-front disincentive to stock market entry.

We find limited evidence that alternative economic or preference-based motives drive continued nonparticipation among pre-lottery nonparticipants. We do, however, find compelling evidence that channels related to beliefs and information play important roles in continued nonparticipation: treatment effects are substantially larger for households that are more educated, are more cognitively-able, win following years of positive equity returns, and experienced above median equity returns between ages 18-25. The credibility of beliefs and information channels as an explanatory factor are also supported by our structural model. We find that the model more accurately predicts estimated effects when we either account for the distribution of subjective equity returns or restrict analysis to a subsample that likely has better informed beliefs.

A parsimonious interpretation of our overall pattern of results is that a substantial fraction of nonparticipating households have a much stronger reluctance to engage with stocks (though not bonds) as an asset class than has been previously documented, and that this reluctance is correlated with factors that indicate less-informed beliefs or higher information processing costs. We cannot rule out all alternative explanations for our results, particularly those that are correlated with proxies for beliefs or information costs. For example, trust is correlated with education (Guiso, Sapienza and Zingales (2004)) and has been previously proposed as an explanation for nonparticipation (Guiso, Sapienza and Zingales (2008)). However, our results provide new insights into the size, structure, and credible sources of economic disincentives that determine stock market participation.

Finally, our paper echoes recent discussions by Kahn and Whited (2016), Lewbel (2016), and Nakamura and Steinsson (2017) by demonstrating the value of both causal and structural estimates and how they can be combined to help interpret economic and financial behavior. Without our causal estimates, the structurally estimated entry costs are significantly understated. And without the structural model, it is impossible to quantify the size and structure of disincentives. Our research design thus demonstrates the benefits of combining causal estimates and identification via economic theory in economic research.

References

Alan, Sule, "Entry costs and stock market participation over the life cycle," *Review of Economic Dynamics*, 2006, *9* (4), 588–611.

Andersen, Steffen and Kasper Meisner Nielsen, "Participation constraints in the stock market: Evidence from unexpected inheritance due to sudden death," *Review of Financial Studies*, 2011, *24* (5), 1667–1697.

Ang, Andrew, Geert Bekaert, and Jun Liu, "Why stocks may disappoint," *Journal of Financial Economics*, 2005, *76* (3), 471–508.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg, "The hyperbolic consumption model: Calibration, simulation, and empirical evaluation," *The Journal of Economic Perspectives*, 2001, *15* (3), 47–68.

Attanasio, Orazio P and Monica Paiella, "Intertemporal consumption choices, transaction costs and limited participation in financial markets: reconciling data and theory," *Journal of Applied Econometrics*, 2011, *26* (2), 322–343.

Barberis, Nicholas, M. Huang, and Richard H Thaler, "Individual preferences, monetary gambles, and stock market participation: A case for narrow framing," *American Economic Review*, 2006, *96* (4), 1069–1090.

Barillas, Francisco and Jesús Fernández-Villaverde, "A generalization of the endogenous grid method," *Journal of Economic Dynamics and Control*, 2007, *31* (8), 2698–2712.

Bazdresch, Santiago, R Jay Kahn, and Toni M Whited, "Estimating and testing dynamic corporate finance models," *The Review of Financial Studies*, 2017, *31* (1), 322–361.

Becker, Thomas A and Reza Shabani, "Outstanding debt and the household portfolio," *Review of Financial Studies*, 2010, *23* (7), 2900–2934.

Benjamin, Daniel J, Sebastian A Brown, and Jesse M Shapiro, "Who is behavioral? Cognitive ability and anomalous preferences," *Journal of the European Economic Association*, 2013, *11* (6), 1231–1255.

Berkelaar, Arjan B, Roy Kouwenberg, and Thierry Post, “Optimal portfolio choice under loss aversion,” *The Review of Economics and Statistics*, 2004, *86* (4), 973–987.

Brunnermeier, Markus K and Stefan Nagel, “Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals’ Asset Allocation,” *American Economic Review*, 2008, *98* (3), 713–736.

Calvet, Laurent, Claire Célérier, Paolo Sodini, and Boris Vallée, “Can Financial Innovation Solve Household Reluctance to Take Risk,” *Working Paper*, 2017.

Calvet, Laurent E and Paolo Sodini, “Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios,” *Journal of Finance*, 2014, *69* (2), 867–906.

_ , John Y Campbell, and Paolo Sodini, “Down or Out: Assessing the Welfare Costs of Household Investment Mistakes,” *Journal of Political Economy*, 2007, *115* (5), 707–747.

Campanale, Claudio, Carolina Fugazza, and Francisco Gomes, “Life-cycle portfolio choice with liquid and illiquid financial assets,” *Journal of Monetary Economics*, 2015, *71*, 67–83.

Campbell, John Y., “Household Finance,” *Journal of Finance*, 2006, *61* (4), 1553–1604.

Carroll, Christopher D, “Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis,” *The Quarterly journal of economics*, 1997, *112* (1), 1–55.

_ , “Portfolios of the Rich,” in “Household Portfolios: Theory and Evidence,” MIT Press, 2002.

_ , “The method of endogenous gridpoints for solving dynamic stochastic optimization problems,” *Economics letters*, 2006, *91* (3), 312–320.

_ and Andrew A Samwick, “The nature of precautionary wealth,” *Journal of monetary Economics*, 1997, *40* (1), 41–71.

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace, “Wealth, health, and child development: Evidence from administrative data on Swedish lottery players,” *The Quarterly Journal of Economics*, 2016, *131* (2), 687–738.

Choi, James J and T Mertens, “Extrapolative expectations and the equity premium,” *Yale University*, 2013.

Cocco, Joao F, “Portfolio choice in the presence of housing,” *Review of Financial Studies*, 2005, *18* (2), 535–567.

_ , Francisco J Gomes, and Pascal J Maenhout, “Consumption and portfolio choice over the life cycle,” *Review of Financial Studies*, 2005, *18* (2), 491–533.

Cooper, Russell and Guozhong Zhu, “Household finance over the life-cycle: What does education contribute?,” *Review of Economic Dynamics*, 2016, *20*, 63–89.

Davis, Steven J, Felix Kubler, and Paul Willen, “Borrowing costs and the demand for equity over the life cycle,” *Review of Economics and Statistics*, 2006, *88* (2), 348–362.

Dimmock, Stephen G and Roy Kouwenberg, “Loss-aversion and household portfolio choice,” *Journal of Empirical Finance*, 2010, *17* (3), 441–459.

Dominitz, Jeff and Charles F Manski, “Measuring and interpreting expectations of equity returns,” *Journal of Applied Econometrics*, 2011, *26* (3), 352–370.

Epstein, L. G. and Stan Zin, “Substitution, Risk Aversion, and the Temporal Behavior of Choice and Asset Returns,” *Journal of Political Economy*, 1991.

Erickson, Timothy and Toni M Whited, “Two-step GMM estimation of the errors-in-variables model using high-order moments,” *Econometric Theory*, 2002, *18* (3), 776–799.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso, “Asset Market Participation and Portfolio Choice over the Life-Cycle,” *The Journal of Finance*, 2017, *72* (2), 705–750.

_ , Luigi Guiso, and Luigi Pistaferri, “Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk,” *The Review of Economic Studies*, 2017, *85* (1), 437–474.

Flavin, Marjorie and Takashi Yamashita, “Owner-occupied housing: life-cycle implications for the household portfolio,” *American Economic Review*, 2011, *101* (3), 609–614.

Fuster, Andreas and Basit Zafar, “The sensitivity of housing demand to financing conditions: evidence from a survey,” *FRB of New York Staff Report*, 2014, (702).

Gomes, Francisco and Alexander Michaelides, “Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence,” *Journal of Finance*, 2005, *60* (2), 869–904.

Gourieroux, Christian, Alain Monfort, and Eric Renault, “Indirect inference,” *Journal of applied econometrics*, 1993, *8* (S1).

Gourinchas, Pierre-Olivier and Jonathan A Parker, “Consumption over the life cycle,” *Econometrica*, 2002, *70* (1), 47–89.

Greenwood, Robin and Andrei Shleifer, “Expectations of returns and expected returns,” *Review of Financial Studies*, 2014, *27* (3), 714–746.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa, “IQ and stock market participation,” *The Journal of Finance*, 2011, *66* (6), 2121–2164.

Grossman, Sanford J and Guy Laroque, “Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods,” *Econometrica*, 1990, *58* (1), 25–51.

Guiso, Luigi and Paolo Sodini, “Household finance: An emerging field,” *Handbook of the Economics of Finance*, 2013, *Part B*, 1397–1532.

—, Michael Haliassos, and Tullio Jappelli, *Household portfolios*, MIT press, 2002.

—, Paola Sapienza, and Luigi Zingales, “The role of social capital in financial development,” *American economic review*, 2004, *94* (3), 526–556.

—, —, and —, “Trusting the stock market,” *Journal of Finance*, 2008, *63* (6), 2557–2600.

Haliassos, Michael and Carol C Bertaut, “Why do so few hold stocks?,” *The Economic Journal*, 1995, *105* (432), 1110–1129.

Hankins, Scott, Mark Hoestra, and Paige Marta Skiba, “The Ticket to Easy Street? The Financial Consequences of Winning the Lottery,” *Review of Economics and Statistics*, August 2011, *93* (3), 961–969.

Heaton, John and Deborah Lucas, “Portfolio choice and asset prices: The importance of entrepreneurial risk,” *Journal of Finance*, 2000, *55* (3), 1163–1198.

— and —, “Portfolio choice in the presence of background risk,” *The Economic Journal*, 2000, *110* (460), 1–26.

Hurd, Michael D, "Subjective probabilities in household surveys," *Annu. Rev. Econ.*, 2009, *1* (1), 543–562.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote, "Estimating the effect of unearned income on labor supply, earnings, savings and consumption: Evidence from a survey of lottery players," *The American Economic Review*, 2001, *91* (4), 778–794.

Kahn, R. Jay and Toni Whited, "Identification with Models and Exogenous Data Variation,," *Foundations and Trends in Accounting,*, 2016, *10*, 361–375.

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

Khorunzhina, Natalia, "Structural estimation of stock market participation costs," *Journal of Economic Dynamics and Control*, 2013, *37* (12), 2928–2942.

Laibson, David, "Golden eggs and hyperbolic discounting," *The Quarterly Journal of Economics*, 1997, pp. 443–477.

Laun, Tobias and Johanna Wallenius, "A life cycle model of health and retirement: The case of Swedish pension reform," *Journal of Public Economics*, 2015, *127*, 127–136.

Lewbel, Arthur, "The Identification Zoo-Meanings of Identification in Econometrics," *Forthcoming on Journal of Economic Literature*, 2016, *7.*

Lindqvist, Erik and Roine Vestman, "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–128.

Love, David and Gregory Phelan, "Hyperbolic discounting and life-cycle portfolio choice," *Journal of Pension Economics & Finance*, 2015, *14* (4), 492–524.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S Mitchell, "Optimal financial knowledge and wealth inequality," *Journal of Political Economy*, 2017, *125* (2), 431–477.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126*, 373–416.

Mankiw, N Gregory and Stephen P Zeldes, "The consumption of stockholders and nonstockholders," *Journal of Financial Economics*, 1991, *29* (1), 97–112.

Mayer, Christopher J and Gary V Engelhardt, "Gifts, down payments, and housing affordability," *Journal of Housing Research*, 1996, *7* (1), 59.

McKay, Alisdair, "Search for financial returns and social security privatization," *Review of Economic Dynamics*, 2013, *16* (2), 253–270.

Merton, Robert C, "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 1971, *3* (4), 373–413.

Mulligan, Casey B and Xavier Sala i Martin, "Extensive margins and the demand for money at low interest rates," *Journal of Political Economy*, 2000, *108* (5), 961–991.

Nakamura, Emi and Jón Steinsson, "Identification in Macroeconomics," Technical Report, National Bureau of Economic Research 2017.

Nekoei, Arash and David Seim, "How do Inheritances Shape Wealth Inequality? Theory and Evidence from Sweden," *Working Paper*, 2018.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Information acquisition and under-diversification," *The Review of Economic Studies*, 2010, *77* (2), 779–805.

Paiella, Monica, "The forgone gains of incomplete portfolios," *Review of Financial Studies*, 2007, *20* (5), 1623–1646.

Poterba, James M and Andrew A Samwick, "Taxation and household portfolio composition: US evidence from the 1980s and 1990s," *Journal of Public Economics*, 2003, *87* (1), 5–38.

Rooij, Maarten CJ Van, Annamaria Lusardi, and Rob JM Alessie, "Financial Literacy, Retirement Planning and Household Wealth," *The Economic Journal*, 2012, *122* (560), 449–478.

Samuelson, Paul A, "Lifetime portfolio selection by dynamic stochastic programming," *Review of Economics and Statistics*, 1969, *51* (3), 239–246.

Smith, Anthony A, "Estimating nonlinear time-series models using simulated vector autoregressions," *Journal of Applied Econometrics*, 1993, *8* (S1).

Tversky, Amos and Daniel Kahneman, "Rational choice and the framing of decisions," *Journal of business*, 1986, pp. S251–S278.

Vestman, Roine, "Limited Stock Market Participation Among Renters and Home Owners," *SSRN No. 1573181*, 2013.

Viceira, Luis M, "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *Journal of Finance*, 2001, *56* (2), 433–470.

Vissing-Jørgensen, Annette, "Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions," *NBER Macroeconomics Annual*, 2003, *18*, 139–208.

Waldenström, Daniel, "Swedish stock and bond returns, 1856–2012," *IFN Working Paper No. 1027*, 2014.

World Health Organization, "Risk Factors," in Margie Peden, Richard Scurfiled, David Sleet, Dinesh Mohan, Adnan A Hyder, and Eva Jarawan, eds., *World report on road traffic injury prevention*, World Health Organization, 2004, chapter 3, pp. 69–106.

A Structural Model Details

A.1 Household Decision Problem

The full household decision problem described in Section 4 is written as:

$$V_t(X_t, P_t, I_t, L_t, e,m) = \max_{C_t, Part_t, \alpha_t} \Bigg\{ (1-\beta s_t) C_t^{1-1/\psi} +$$

$$\beta \mathbb{E}\left[s_t V_{t+1}(X_{t+1}, P_{t+1}, I_{t+1}, e, m)^{1-\rho} + (1-s_t) b(X_{t+1})^{1-1/\psi} \right]^{\frac{1-1/\psi}{1-\rho}} \Bigg\}^{\frac{1}{1-1/\psi}}$$

$$X_{t+1} = \left[R_f + \alpha_t (R^s_{t+1} - R_f)\right](X_t - C_t) + Y_{t+1} - \left[(1-I_t)\times \chi + \kappa)\right] \times Part_t + L_t$$

$$0 \leq \alpha \leq 1$$

$$Y_t = \begin{cases} exp(f(t, Z_t)) P_t U_t \;\; if \;\; t \leq t_R \\ \lambda_{e,m} Y_{t_R} \;\; if \;\; t > t_R \end{cases}$$

$$P_t = P_{t-1} N_t$$

$$I_{t+1} = (1 - I_t) \times Part_t$$

$$\begin{pmatrix} r^s_t - r_f \\ log(N_t) \\ log(U_t) \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mu_s \\ -\sigma^2_N/2 \\ -\sigma^2_U/2 \end{pmatrix}, \begin{pmatrix} \sigma^2_s & \rho_{n,r} \times \sigma_n \sigma_s & 0 \\ \rho_{n,r} \times \sigma_n \sigma_s & \sigma^2_n & 0 \\ 0 & 0 & \sigma^2_u \end{pmatrix} \right]$$

A.2 Estimation and Test Statistics

We consider two test statistics to check for overidentifying restrictions and to evaluate model fit. First the standard overidentifying test used to test the model's fit of the empirical moments, correcting for simulation error, is given by

$$\frac{nK}{1+K} g(\nu_{i,s}, \theta)' \hat{\Omega}^{-1} g(\nu_{i,s}, \theta) \rightarrow \chi^2_{|g(\nu_{i,s},\theta)| - |\theta|}. \tag{13}$$

Second, we consider the Wald test for external validity presented in Bazdresch et al. (2017) that considers the model's fit of non-targeted moments $m^\star$. The null hypothesis of non-targeted fit,

$$g^\star(\nu_{i,s}, \theta) = \mathbb{E}\left[m^\star(\nu_i) - \frac{1}{K}\sum_{k=1}^{K} m^\star(\nu_{i,s}^k(\theta))\right] = 0, \tag{14}$$

can be tested by a Wald statistic defined as

$$g^\star(\nu_{i,s}, \hat{\theta})' avar(g^\star(\nu_{i,s}, \hat{\theta}))^{-1} g^\star(\nu_{i,s}, \hat{\theta}) \to \chi^2_{|g^\star(\nu_{i,s}, \hat{\theta})|} \tag{15}$$

$$avar(g^\star(\nu_{i,s}, \hat{\theta})) = \mathbb{E}\left[\phi_g^\star \phi_g^\star\right] \tag{16}$$

where $\phi_g^\star$ denotes the influence function for $g^\star$.

A.3 Model Solution

To solve the model, we follow Carroll (1997) and normalize the value function, state variables, and controls by the permanent component of income P_t to eliminate P_t as a state variable. We use lower case letters to denote the normalized variables (e.g., $v_t = V_t/P_t$, $x_t = X_t/P_t$). After these transformations, the model is solved by backwards induction. We assume that the last period's utility is as $v_T = b(x_T)^{1-\psi}$. We then use this to solve for the optimal saving policy $x_{T-1} - c_{T-1}$ using the endogenous grid method and portfolio allocation α_{t-1} using a grid search (100 grid points) (Carroll (2006); Barillas and Fernández-Villaverde (2007)). For points that do not fall on next period's stored state-grid, we use cubic interpolation to evaluate the value function. To calculate the expected value of next period's value function, we follow the procedure described in Gomes and Michaelides (2005) to create a state transition matrix that makes integration less computationally costly. After having obtained the optimal saving and portfolio allocation policies, we are able to calculate the v_{T-1} value function. We then repeat this process and iterate backwards until reaching age t_0. We repeat this for all combinations of marital status and education level, and store the resulting policy functions.

A.4 Survival Probability Estimation

The survival probability is calculated using the observed survival probabilities from years 1999-2000. We select 100,000 individuals in year 1998 from the Swedish population, and define a binary indicator equal to one if the individual is observed alive in 1999. We then regress a quartic in age on this indicator. We do not permit time or cohort effects in our estimation, and do not allow survival probabilities to vary with wealth, income, or sex. There is no attrition or selection concerns in this sample as it is drawn randomly from the entire population. The resulting estimates are presented in Figure A.1.

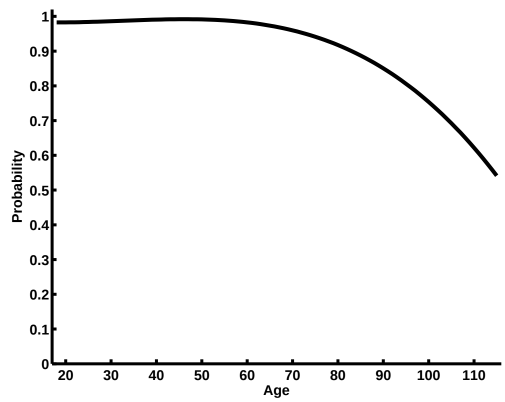


Figure A.1: Survival Probabilities. This figure presents the estimated one year survival probability for each age.

A.5 Income Estimation

Our estimation of income profiles follows the procedure described in Cocco et al. (2005). Our definition of income is total income after taxes and transfers. As noted in Cocco et al. (2005), because there are (potentially endogenous) insurance mechanisms – including government transfers, family transfers, and spousal labor supply decisions – that provide a lower bound on income (perhaps especially in countries with strong social safety nets such as Sweden) this definition captures this insurance without explicit modeling of all income smoothing mechanisms. Our estimation sample is the sample of lottery winners in the thirty years (or as many as possible) prior to the lottery event.

Income processes are estimated separately for each of the education groups we consider. Our estimation sample is our sample of lottery winners prior to the lottery year. We regress the log of income on dummies of age and marital status. We then regress a third-order polynomial in age on the age dummies and marital status for housholds between ages 18-65 to recover an average income profile $f(t, m, e)$. The resulting average income profile estimates $exp(f(t, m, e))$ are shown in Figure A.2, with dotted lines representing married households and dashed lines representing single households. $P_{i,s}$ is then constructed as the ratio of observed to average income for each household in our sample.

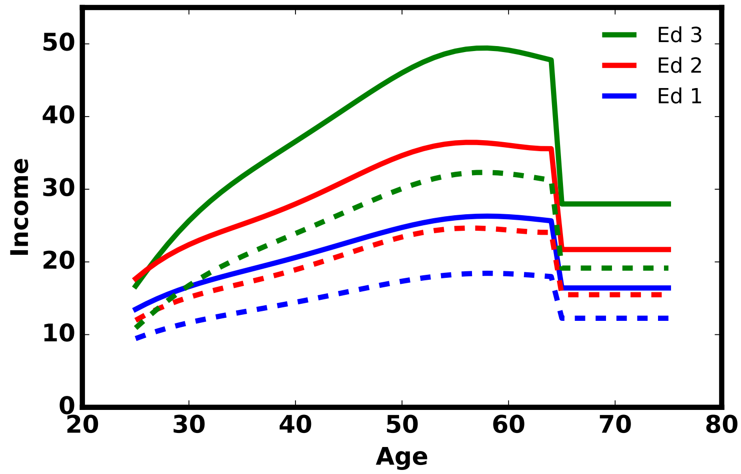


Figure A.2: Average Income Profiles. This figure presents the deterministic income component $f(t, m, e)$. Solid lines reflect married households while dashed lines reflect single households.

We estimate income variance parameters again following Cocco et al. (2005), who closely follow the procedure proposed by Carroll and Samwick (1997). In particular, defining

$$\epsilon^Y_{i,t} \equiv log(Y_{i,t}) - \hat{f}(t, m_i, e_i)$$
$$r_{i,d} \equiv \epsilon^Y_{i,t+d} - \epsilon^Y_{i,t},$$

then because

$$\text{Var}(r_{i,d}) = d\sigma^2_N + 2 * \sigma_U$$

we can recover $\sigma_{N,e}$ and $\sigma_{U,e}$ via OLS regresison on Var$(r_{i,d})$ on d for each separate education group.

To estimate the correlation between income and equity returns, note that $\epsilon^Y_{i,t}$ can be written as

$$r_{i,1} = log(i, N_t) + log(U_{i,t}) - log(U_{i,t+1})$$

and taking the average yields

$$r_{i,1} = log(N_{i,t}) + log(U_{i,t}) - log(U_{i,t+1})$$

Decomposing $N_{i,t}$ into aggregate and idiosynchratic components, letting s index year, and averaging (for each education group) yields:

$$\bar{r}_{i,1,s,e} = log(N_{s,e}^{Agg}).$$

The correlation between equity returns and $log(N_{i,t})$ for each education group is then recovered by the coefficient from an OLS regression of $\bar{r}_{i,1,s}$ on excess returns, where excess returns are defined as the difference between Stockholm Stock Exchange and short-term Swedish Treasury returns (Waldenström (2014)).

A.6 Retirement Income Replacement Rates

Retirement income replacement rates are approximated using the formulas described in Section 3 of Laun and Wallenius (2015), which conducts a detailed analysis of the Swedish pension system. Our formulas are slightly simplified due to the assumption that labor supply is exogenous. The pension has two parts. First, all households receive 96% of a basic amount (BA) of 43,600 SEK (6,500 USD). Second, an earning supplement is given by

$$.6 \times AP \times BA$$

where AP denotes pension points calculated from the fifteen years with highest observed income calculated recursively by the following formula:

$$AP_{t+1} = AP_t + \frac{1}{15} \max\left(0, \frac{\min(Y_t, 7.5BA) - BA}{BA} - AP_t\right)$$

Thus, retirement income is approximated as the ratio of the following formula

$$.6 \times AP \times BA + .96BA$$

to age 65 income.

To conserve state variables, we do not carry pension points as a state variable as in Laun and

Wallenius (2015). Instead, we simulate 20,000 income processes for each education and marital status, and calculate the average replacement rate for each group.

A.7 Model Benchmarks and Fit

Below we present the full specification of the regressions that form our EPF benchmarks. In addition, we indicate the corresponding panel for each regression in Table A.1, and, when appropriate, the location of selected coefficients presented in Table 6. Empirical estimates are presented in A.1, Column(1). Note in all lottery regressions we include cell-fixed effects that ensure all identifying variation comes from players in the same cell. The regressions we consider are:

1. Pre-lottery regressions (Table A.1, Panel A.i-ii):

$$\begin{aligned} c_{i,s} &= b + b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_I I_{i,s} + \eta_{i,s}^C \\ Part_{i,s} &= b + b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_I I_{i,s} + \eta_{i,s}^{Part} \end{aligned} \tag{17}$$

2. Post-lottery regressions (Table A.1, Panel B.i-ii; Table 6, Panel B.i):

$$\begin{aligned} c_{i,s} &= b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_I I_{i,s} + b_l l_{i,s} + MX_{i,0} + \eta_{i,s}^C \\ Part_{i,s} &= b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_I I_{i,s} + b_l l_{i,s} + MX_{i,0} + \eta_{i,s}^{Part} \end{aligned} \tag{18}$$

3. Post-lottery regressions by participation status (Table A.1; Panel B.iii-iv, Table 6, Panel B.ii)). These regressions are estimated separately in subsamples restricted fo participants $I_{i,s} = 1$ and nonparticipants ($I_{i,s} = 0$):

$$\begin{aligned} c_{i,s} &= b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_l l_{i,s} + MX_{i,0} + \eta_{i,s}^C \\ Part_{i,s} &= b_t t_{i,s} + b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + b_l l_{i,s} + MX_{i,0} + \eta_{i,s}^{Part} \end{aligned} \tag{19}$$

4. Post-lottery regressions, nonlinear (Table A.1; Panel B.v, Table 6, Panel B.iii):

$$\begin{aligned} Part_{i,s} = b_t t_{i,s} &+ b_{t^2} t_{i,s}^2 + b_x x_{i,s} + b_{x^2} x_{i,s}^2 + \mathbb{1}_{l_{i,s} \in [1.5,15)} + \\ &\mathbb{1}_{l_{i,s} \in [15,150)} + \mathbb{1}_{l_{i,s} \in [150,300)} + \mathbb{1}_{l_{i,s} \in [300,\infty)} + MX_{i,0} + \eta_{i,s}^{Part} \end{aligned} \tag{20}$$

Table A.1 presents the fits of are various estimation exercises. Our pre-lottery estimation (Table

Table A.1: Structural Estimation Model Fit. This table presents the model fit for our various structural estimation exercises. Column (1) presents the empirical estimate for each regression, while Column (2)-(6) presents the model-predictions for the indicated estimation. Only matched coefficients are included in this table, and omitted coefficients are not targetted. All estimations use the post-1999 sample of lottery winners.

	Empirical Estimate (1)	Pre-Lottery (2)	Post-Lottery (3)	Pre-/Post-Lottery Combined (4)	Pre- and Post-lottery: Nonlinear (5)	Pre-lottery, Present-biased preferences (6)
A. Pre-Lottery Benchmarks						
i. Consumption						
Age	.619	.077		.185		.164
Age^2	-.006	-.001		-.002		-.002
$Wealth/PI$	.164	.203		.156		.173
$(Wealth/PI)^2$	.000	.000		.000		.000
$Part_{s-1}$	-.881	1.388		1.435		1.797
$Constant$	4.508	7.860		6.993		7.644
ii. Participation						
Age	.000	.000		.000		.003
Age^2	.000	.000		.000		.000
$Wealth/PI$	.000	.000		.000		.000
$(Wealth/PI)^2$	.000	.000		.000		.000
$Part_{s-1}$	.883	.927		.938		.935
$Constant$	.114	.030		.019		-.013
B. Lottery Benchmarks						
i. Consumption						
Age	.614		.250	.223		
Age^2	-.005		-.003	-.003		
$Wealth/PI$	.039		.139	.155		
$(Wealth/PI)^2$	.000		.000	.000		
$Part_{s-1}$	-1.618		1.487	1.253		
$Lottery$	.185		.123	138		
ii. Participation						
Age	.001		.000	.001		
Age^2	.000		.000	.000		
$Wealth/PI$	.000		.000	.000		
$(Wealth/PI)^2$	.000		.000	.000		
$Part_{s-1}$	.796		.993	.933		
$Lottery$	.028		.030	.067	.029	
iii. Effect on Consumption by Prior Participation Status						
$Lottery$, Nonparticipants	.239		.121	.136		
$Lottery$, Participants	.166		.124	.138		
iv. Effect on Participation by Prior Participation Status						
$Lottery/1M\ SEK$, Nonparticipants	.104		.137	.292	.104	
$Lottery/1M\ SEK$, Participants	.002		.000	.000	.000	
v. Effect on Participation by Prize Size (USD), Nonparticipants						
$1.5K < L_i \leq 15K$	-.012				.006	
$15K < L_i \leq 150K$	.078				.080	
$150K < L_i \leq 300K$	.156				.158	
$300K < L_i$	.359				.357	

6, Column (1)), which targets only pre-lottery regressions (Table A.1, Panel A.i and Panel A.ii) is presented in Column (2). Our post-lottery estimation (Table 6, Column (2)), which targets only post-lottery regressions (Table A.1, Panel B.i-iv) is presented in Column (3). Our pre-/post-lottery combined estimation (Table 6, Column (3)), which targets only both pre- and post-lottery regressions (Table A.1, Panels A.i-ii, B.i-iv) is presented in Column (4). Our entry-cost heterogeneity estimation (Table 6, Column (4)), which targets selected post-lottery regression coefficients of the effect of lottery prizes on participation (Table A.1, Panel B.ii, iv-v) is presented in Column (5). Finally, our estimation with q present-biased preferences (Table 8, Column (2)), which targets only pre-lottery regressions (Table A.1, Panel A.i and Panel A.ii)) is presented in Column (6).

A.8 Lifecycle Profiles Comparison

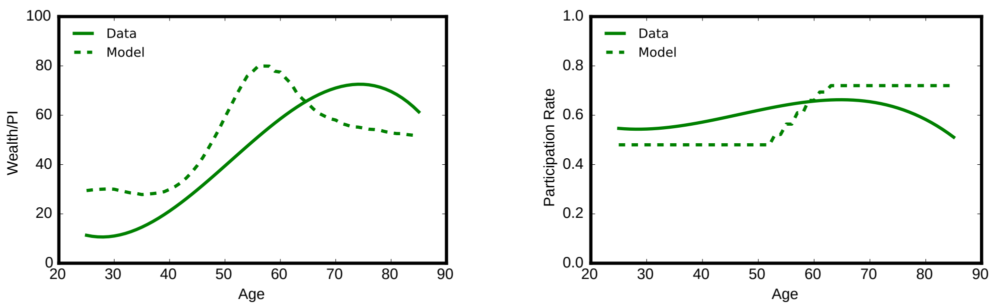
In this section we compare the lifecycle profiles implied by our model estimates to their empirical counterparts. To estimate empirical lifecycle profiles of stock market participation and wealth, we use a simplified version of the estimation procedure described in Fagereng et al. (2017a). Our estimation sample in this exercise consists of the matched population sample presented in Table 4, Column (2).

To estimate lifecycle profiles of the wealth/income ratio, we run an OLS regression of the registry defined wealth/income ratio on age indicators, year indicators, and a proxy of cohort effects defined by the average returns on the Stockholm Stock Exchange experienced between ages 18-25. We then regress the predicted wealth-to income ratios for each age on a cubic polynomical of age. The resulting wealth-to-income profiles are presented as the dotted line in panel (a) of the below figures.

To estimate lifecycle profiles of stock market participation, we run a probit regression of household stock market participation on age indicators, year indicators, and a proxy of cohort effects defined by the average returns on the Stockholm Stock Exchange experienced between ages 18-25. We then regress the predicted stock market participation probabilities for each age on a cubic polynomical of age. The resulting participation probabilities profiles are presented as the dotted line in panel (b) of the below figures. Overall, our estimated wealth and participation profiles are similar to those obtained by Fagereng et al. (2017a) for a representative Norwegian sample.

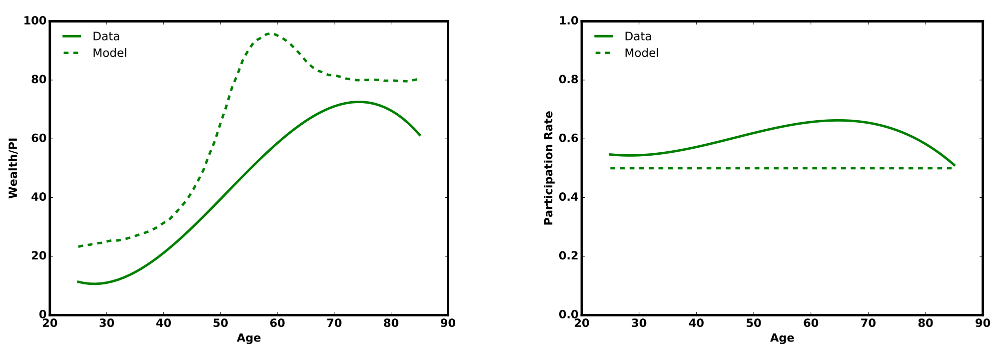
To generate model implied profiles, we draw a random sample of 10000 Swedish households aged 18-25 between 1999-2004. Because marital and education histories are incomplete by this age, we assign marital and education status as the highest values observed by 2009. We then record all model state variables, and simulate saving and participation decisions through age 85. The average of these simulations for each age are presented as the dotted lines in the below figures.

Figure A.3 presents results from our model using parameter estimates from our estimation with pre-lottery data (Table 6, Column (1)). Figure A.4 presents results from our model using parameter estimates from our estimation with post-lottery data (Table 6, Column (2)). Figure A.5 presents results from our model using parameter estimates from our estimation with pre- and post-lottery data (Table 6, Column (3)). Figure A.6 presents results from our model with quasi-hyperbolic discounting using parameter estimates from our estimation with pre- and post-lottery data (Table 8, Column (2)).



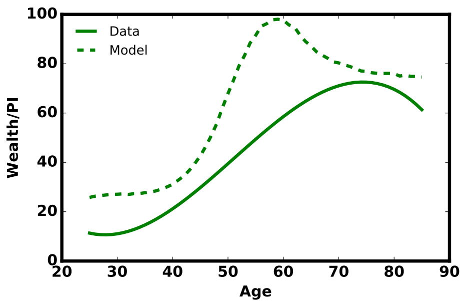
(a) Wealth/PI ratio **(b) Participation**

Figure A.3: Lifecycle profiles - pre-lottery data. This figure compares the model-predicted and empricial wealth/permanent income ratio and participation rate over the lifecycle. The model is simulated using estimates obtained from pre-lottery data (Table 6, Column (1))

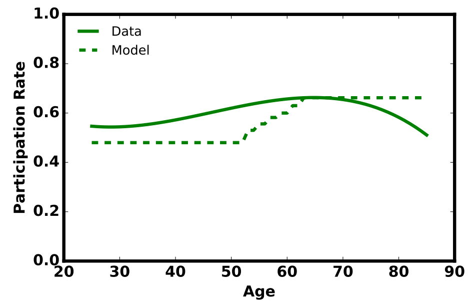


(a) Wealth/PI ratio **(b) Participation**

Figure A.4: Lifecycle profiles - lottery data. This figure compares the model-predicted and empricial wealth/permanent income ratio and participation rate over the lifecycle. The model is simulated using estimates obtained from lottery data (Table 6, Column (2))

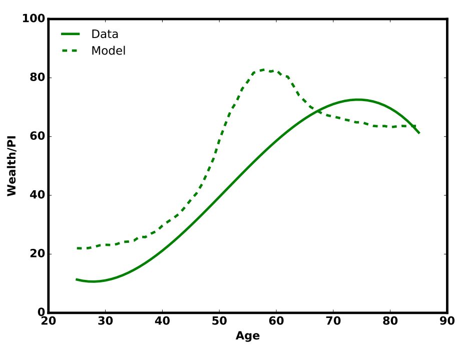


(a) Wealth/PI ratio

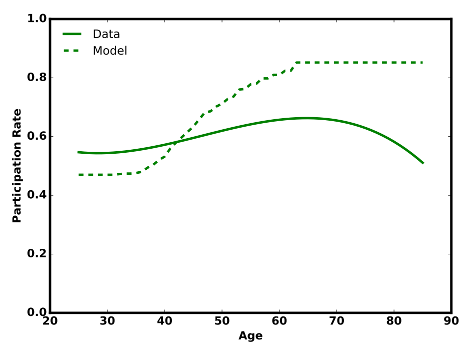


(b) Participation

Figure A.5: Lifecycle profiles - pre- and post-lottery. This figure compares the model-predicted and empricial wealth/permanent income ratio and participation rate over the lifecycle. The model is simulated using estimates obtained from pre- and post-lottery data (Table 6, Column (3))



(a) Wealth/PI ratio



(b) Participation

Figure A.6: Lifecycle Profiles - Pre-lottery with quasi-hyperbolic discounting. This figure compares the model-predicted and empricial wealth/permanent income ratio and participation rate over the lifecycle for our model that allows for $\beta - \delta$ preferences. Parameters are obtained by setting $\beta = .6$ and re-estimating the model with pre-lottery data (Table 8, Column (2))

B Online Appendix - Supplemental Tables

Table B.1: Summary Statistics and Prize Distributions for Triss Lump-sum and Monthly Prize Winners.

Summary Statistics

	A. All-Year		**B. Post-1999**	
Prize Amount	**Triss Lump-sum**	**Triss Monthly**	**Triss Lump-sum**	**Triss Monthly**
Female	0.54	0.49	0.56	0.48
Age (years)	52.52	49.78	51.86	50.50
Nordic Born	0.95	0.94	0.94	0.94
Household Members (#)	0.52	0.54	0.59	0.56
Household Income (K SEK)	335	374	382	392
Married	0.52	0.54	0.54	0.55
Higher Education	0.24	0.26	0.26	0.28
Financial				
Net Wealth (K SEK)			857	736
Gross Debt (K SEK)			448	386
Home Owner			0.73	0.66
Equity Owner			0.63	0.59
N	3,399	476	1,776	386

Prize Distribution

	A. All-Year		**B. Post-1999**	
Prize Amount	**Triss Lump-sum**	**Triss Monthly**	**Triss Lump-sum**	**Triss Monthly**
$L_i = 0$	0	0	0	0
$L_i \leq 10K$	0	0	0	0
$10K < L_i \leq 100K$	985	0	366	0
$100K < L_i \leq 500K$	2074	0	1237	0
$500K < L_i \leq 1M$	157	0	89	0
$1M < L_i \leq 2M$	49	130	22	110
$2M < L_i$	134	346	62	276
Total	3,399	476	1,776	386

Table B.2: Demographic and Financial Predictors of Participation in Post-1999 Sample and Sex- and Age-Weighted Swedish Representative Sample. The regression model is estimated using year-end net wealth in 1999 and is comparable to that used by Calvet et al. (2007). Marginal effects are calculated as the predicted effect of a one-standard deviation change on the probability of participation, holding fixed the value of all other variables at their median value.

	Post-1999 Lottery			**Matched Population**		
	Estimate	**SE**	**Marginal Effects.**	**Estimate**	**SE**	**Marginal Effects**
	(1)	(2)	(3)	(4)	(5)	(6)
Income	0.064	0.006	5.43%	0.057	0.008	4.87%
Financial Assets	0.190	0.005	68.68%	0.177	0.007	76.58%
Total Real Estate	0.028	0.006	14.09%	0.036	0.011	17.64%
Total Liabilities	-0.014	0.007	-6.91%	-0.023	0.011	-11.14%
Retired	0.010	0.029	0.17%	0.056	0.045	0.76%
Self-Employed	0.046	0.043	0.35%	0.039	0.059	0.30%
Unemployed	0.020	0.028	0.24%	0.068	0.042	0.88%
Student	0.135	0.047	1.09%	0.068	0.055	0.87%
Age	-0.013	0.001	-7.74%	-0.013	0.001	-8.65%
Household Size	-0.044	0.012	-1.34%	-0.050	0.015	-1.88%
High School	0.145	0.024	2.77%	0.176	0.039	3.13%
Higher Degree	0.224	0.026	4.28%	0.294	0.042	5.34%
Missing Education	0.111	0.107	0.44%	0.421	0.127	1.52%
Immigrant	-0.166	0.052	-1.11%	-0.262	0.059	-2.71%
Constant	-3.884	0.123	.	-3.594	0.172	.
N	70,166			70,166		

Table B.3: Effect of Wealth (1M SEK) on Participation Probability. This table presents coefficients, standard errors, sample size, and mean predicted participation probability when lottery wealth is zero ($\hat{y}|L_i = 0$) obtained from estimating Equation 1 in the all-year sample. Columns 5 through 8 show analogous estimates with participation defined more narrowly to only include directly owned stocks.

Horizon	**A. Stock or Mutual Fund**				**B. Stock Only**			
(s)	β_s	**SE**	N	$\hat{y}_{\|L_i=0}$	β_s	**SE**	N	$\hat{y}_{\|L_i=0}$
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
-1	0.005	0.012	70,166	0.725	0.002	0.012	70,166	0.425
0	0.039	0.010	75,773	0.727	0.021	0.011	75,773	0.433
1	0.047	0.010	91,940	0.729	0.036	0.011	91,940	0.434
2	0.045	0.010	113,879	0.749	0.041	0.011	113,879	0.462
3	0.043	0.010	141,878	0.761	0.043	0.011	141,878	0.478
4	0.042	0.010	149,324	0.770	0.041	0.012	149,324	0.494
5	0.046	0.010	153,464	0.773	0.033	0.012	153,464	0.502
6	0.038	0.010	168,061	0.778	0.039	0.012	168,061	0.512
7	0.035	0.010	182,380	0.788	0.044	0.012	182,380	0.523
8	0.053	0.010	197,045	0.794	0.046	0.012	197,045	0.533
9	0.049	0.010	216,787	0.794	0.043	0.012	216,787	0.537
10	0.041	0.010	214,208	0.797	0.023	0.012	214,208	0.545

Table B.4: Effect of Wealth (1M SEK) on Participation Probability by $s = -1$ Participation Status. This table presents coefficients, standard errors, sample size, and mean predicted participation probability when lottery wealth is zero ($\hat{y}|L_i = 0$) obtained from estimating Equation 1 in the post-1999 sample stratified by participation status.

Horizon	**Participants**				**Nonparticipants**			
(s)	β_s	**SE**	N	$\hat{y}_{\|L_i=0}$	β_s	**SE**	N	$\hat{y}_{\|L_i=0}$
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
0	0.000	0.005	50,861	0.979	0.120	0.024	19,278	0.070
1	0.002	0.009	47,380	0.961	0.116	0.029	17,276	0.093
2	0.010	0.009	43,487	0.947	0.093	0.032	15,316	0.114
3	0.021	0.007	40,324	0.931	0.073	0.032	13,757	0.129
4	0.018	0.010	36,842	0.917	0.080	0.043	12,267	0.141

Table B.5: Effect of Wealth on Equity Market Participation Probability by Prize Size. Coefficients are obtained by estimating Equation 1 in the post-1999 sample with the lottery wealth variable replaced by indicators for five mutually exclusive prize categories: 0 to 10K (omitted category), 10K to 100K, 100K to 1M, 1M to 2M, and 2M+ SEK. Marginal effects are calculated by dividing the effect-size estimate by the mean prize in each category.

A. Participants

	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate	-0.011	-0.005	0.001	-0.024
SE	0.009	0.016	0.011	0.033
ME	-0.364	-0.026	0.001	-0.007
N	478	801	203	50

B. Nonparticipants

	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate	0.014	0.082	0.177	0.399
SE	0.029	0.037	0.044	0.094
ME	0.382	0.434	0.159	0.127
N	256	525	94	28

Table B.6: Effect of Wealth (1M SEK) on Participation Probability by Form of Payment. This table presents coefficients, standard errors, sample size, and mean predicted participation probability when lottery wealth is zero ($\hat{y}|L_i = 0$) obtained from estimating Equation 1 in the post-1999 sample of Triss-monthly winners stratified by participation status. Columns 1 through 8 show estimates for nonparticipants at time $s = -1$ that received annual and lump-sum payments, respectively. Columns 9 through 16 show analogous estimates for participants at time $s = -1$.

	A. Non-participants								**B. Participants**							
Horizon	**Annual Prizes**				**Lump-Sum Prizes**				**Annual Prizes**				**Lump-Sum Prizes**			
(s)	β_s	**SE**	N	$\hat{y}_{\vert L_i=0}$	β_s	**SE**	N	$\hat{y}_{\vert L_i=0}$	β_s	**SE**	**N**	$\hat{y}_{\vert L_i=0}$	β_s	**SE**	**N**	$\hat{y}_{\vert L_i=0}$
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
0	0.011	0.023	164	0.115	0.105	0.029	710	0.083	0.003	0.012	222	0.920	-0.003	0.007	1066	0.960
1	0.007	0.029	135	0.159	0.079	0.034	591	0.133	0.012	0.015	200	0.854	-0.002	0.013	906	0.917
2	-0.008	0.031	112	0.183	0.050	0.036	482	0.177	0.034	0.014	163	0.790	0.015	0.011	733	0.891
3	-0.036	0.039	92	0.354	0.036	0.037	395	0.167	0.041	0.020	134	0.705	0.030	0.007	600	0.880
4	-0.002	0.050	73	0.267	0.022	0.041	314	0.185	0.030	0.026	107	0.727	0.028	0.008	473	0.871

Table B.7: Effect of Wealth (1M SEK) on Participation Probability, Robustness Checks. This table presents the effect of each 1M SEK wealth on participationat $s = 0$ by lottery in the pooled all-year sample and the post-1999 sample stratified by participants (P) and nonparticipants (NP). Capital Insurance: capital insurance ownership is included in participation definition. Structured Product: retail structured products are included in the ownership definition. Individual Analysis: spousal ownership of equities excluded from participation definition. Probit: marginal effects from Probit instead of OLS. Small Prizes Excluded: prizes of size <50K SEK are dropped from analysis. Large Prizes Excluded: prizes greater than 1.5M SEK are dropped from analysis. For each case, the coefficient from the pooled regression estimated in the all-year sample and the coefficients stratified by pre-lottery participation status estimated in the post-1999 sample are presented.

				Lottery and Subsamples					
	Kombi			**Triss**					
	Pooled	**P**	**NP**	**Pooled**	**P**	**NP**			
Effect	0.042	0.005	0.151	0.036	-0.003	0.105			
SE	0.018	0.007	0.036	0.012	0.007	0.029			
N	28,571	19,154	7,261	1,968	1,066	710			
				Participation Definitions					
	Capital Insurance			**Structured Product**			**Individual Analysis**		
	Pooled	**P**	**NP**	**Pooled**	**P**	**NP**	**Pooled**	**P**	**NP**
Effect	0.052	0.000	0.151	.055	-0.000	0.184	0.043	0.005	0.099
SE	0.010	0.005	0.029	0.010	0.005	0.029	0.011	0.004	0.021
N	75,773	52,339	17,800	75,775	51,425	18,714	75,773	46,703	23,436
				Other Robustness					
	Probit			**Small Prizes Excluded**			**Large Prizes Excluded**		
	Pooled	**P**	**NP**	**Pooled**	**P**	**NP**	**Pooled**	**P**	**NP**
Effect	0.133	0.009	0.532	0.036	-.003	.105	0.050	-.001	.170
SE	0.038	0.072	0.097	.013	.007	.031	0.021	.010	.038
Marginal Effect	0.040	0.000	0.067						
N	75,769	46,918	17,149	2,284	1,267	800	75,680	50,805	19,248

Table B.8: Heterogeneous Effect of Wealth (1M SEK) on Participation Probability among $s = -1$ Equity Market Participants. Coefficients are obtained by estimating Equation 1 at time $s = 0$ in the post-1999 sample of participants at time $s = -1$. Hetero p obtained from an F-test of the null hypothesis that the two lottery-wealth coefficients are identical. Equity returns are based on the MSCI Sweden Index the calendar year prior to the lottery.

A. Rational Channels

	Liquid Wealth		**Home Owner**		**Gross Debt**		**Self-Employed**	
	Low (1)	**High** (2)	**No** (3)	**Yes** (4)	**=0** (5)	**>0** (6)	**No** (7)	**Yes** (8)
Effect	-.001	.000	.001	-.001	.000	-.001	.001	-.012
SE	.006	.008	.014	.005	.006	.006	.005	.017
p	.885	.971	.954	.873	.939	.869	.888	.466
Hetero p	.909		.912		.863		.460	
N	26,585	24,276	9,980	40,881	20,206	30,655	47,631	3,230
$\% Part_{s-1}$	.795	.726	.554	.784	.679	.759	.719	.832

	Net Wealth		**Age**	
	Low (9)	**High** (10)	$\leq$ **61** (11)	**(> 61)** (12)
Effect	-.005	.005	-.001	.001
SE	.009	.002	.007	.007
p	.609	.003	.896	.932
Hetero p	.278		.879	
N	23,701	27,160	26,833	24,028
$\% Part_{s-1}$	.601	.884	.760	.690

B. Information/Beliefs

	Education		**Cognitive Skill**		**Recent Equity Returns**		**Early Equity Returns**	
	No College (15)	**College** (16)	**Low** (17)	**High** (18)	**Low** (19)	**High** (20)	**Low** (21)	**High** (22)
Effect	.002	-.005	-.030	-.032	-.006	.003	-.001	.003
SE	.005	.013	.047	.059	.008	.006	.008	.005
p	.649	.662	.517	.585	.458	.646	.921	.633
Hetero p	.566		.980		.386		.723	
N	36,065	14,796	1,689	3,598	29,849	21,012	27,330	23,531
$\% Part_{s-1}$	.686	.842	.677	.790	.742	.703	.721	.730

Table B.9: Effect of Wealth on Bank Account Balances, Real Assets, Debt, and Structured Products by $s = -1$ Equity Market Participation Status. This table presents coefficients, standard errors, sample size, and mean predicted change in wealth when lottery wealth is zero ($\hat{y}|L_i = 0$) obtained from estimating Equation 1 in the post-1999 sample, stratified by equity market participation status at $s = -1$. The coefficients are interpreted as the effect of 1 SEK – or equivalently the share of the amount won invested in – each asset category. Columns 1-8 show estimates for the effect of wealth on bank account balances, columns 9-16 show the estimates for real estate market, columns 17-24 show the estimates for debt, while columns 25-32 show the estimates for structured products.

	A. Bank								B. Real Estate							
Horizon	Participants				Nonparticipants				Participants				Nonparticipants			
(s)	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
0	0.241	0.040	50,861	.215	0.197	0.035	19,278	.147	-0.032	0.046	50,861	1.185	0.045	0.018	19,278	0.498
1	0.121	0.029	47,380	0.222	0.091	0.021	17,276	0.150	-0.021	0.061	47,380	1.277	0.028	0.019	17,276	0.517
2	0.781	0.024	43,487	0.227	0.055	0.017	15,316	.151	0.013	0.061	43,487	1.379	0.028	0.024	15,316	0.552
3	0.088	0.278	40,324	0.248	0.039	0.017	13,757	0.158	0.062	0.058	40,324	1.452	0.058	0.030	13,757	0.599
4	0.116	0.045	36,842	0.288	0.021	0.013	12,267	0.177	0.031	0.065	36,842	1.560	0.103	0.047	12,267	0.654

	C. Debt								D. Structured Product							
Horizon	Yes				No				Participants				Nonparticipants			
(s)	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$	β_s	SE	N	$\hat{y}_{\|L_i=0}$
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
0	-0.041	0.011	50,861	0.245	-0.031	0.014	19,278	0.160	0.039	0.008	50,862	0.012	0.037	0.011	19,278	0.002
1	-0.056	0.010	47,380	0.247	-0.026	0.017	17,276	0.159	0.049	0.012	47,382	0.013	0.047	0.013	17,276	0.003
2	-0.028	0.017	43,487	0.250	-0.004	0.019	15,316	0.159	0.050	0.010	43,489	0.016	0.030	0.007	15,316	0.003
3	-0.018	0.021	40,324	0.260	0.022	0.020	13,757	0.161	0.045	0.010	40,328	0.020	0.032	0.012	13,757	0.004
4	-0.024	0.017	36,842	0.267	0.044	0.025	12,267	0.165	0.025	0.007	36,848	0.027	0.024	0.011	12,267	0.006

Table B.10: Effect of Wealth (1M SEK) on Participation Probability in Bond, Real Estate, Debt, and Structured Product Markets by $s = -1$ Equity Market Participation Status. This table presents coefficients, standard errors, sample size, and mean predicted participation probability when lottery wealth is zero ($\hat{y}|L_i = 0$) obtained from estimating Equation 1 in the post-1999 sample, stratified by equity market participation status at $s = -1$. Columns 1-8 show estimates for the effect of wealth on bond market and structured product participation, columns 9-16 show the estimates for real estate market participation, columns 17-24 show the estimates for being debt free, while columns 25-32 show the estimates for structured product market participation.

A. Bond

Horizon	Participants				Nonparticipants			
(s)	β_s	SE	N	$\hat{y}_{\vert L_i=0}$	β_s	SE	N	$\hat{y}_{\vert L_i=0}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	0.123	0.017	50,861	0.509	0.173	0.031	19,278	0.225
1	0.149	0.017	47,380	0.532	0.239	0.038	17,276	0.243
2	0.135	0.017	43,487	0.556	0.227	0.040	15,316	0.269
3	0.125	0.021	40,324	0.581	0.188	0.050	13,757	0.288
4	0.111	0.029	36,842	0.605	0.148	0.056	12,267	0.300

B. Real Estate

Horizon	Participants				Nonparticipants			
(s)	β_s	SE	N	$\hat{y}_{\vert L_i=0}$	β_s	SE	N	$\hat{y}_{\vert L_i=0}$
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
0	-0.003	0.008	50,861	0.802	0.028	0.014	19,278	0.578
1	0.003	0.010	47,380	0.801	0.052	0.018	17,276	0.573
2	-0.005	0.014	43,487	0.799	0.036	0.018	15,316	0.574
3	0.012	0.010	40,324	0.800	0.052	0.022	13,757	0.578
4	0.007	0.010	36,842	0.798	0.072	0.032	12,267	0.580

C. Debt-Free

Horizon	Yes				No			
(s)	β_s	SE	N	$\hat{y}_{\vert L_i=0}$	β_s	SE	N	$\hat{y}_{\vert L_i=0}$
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
0	0.048	0.014	50,861	0.401	0.104	0.020	19,278	0.500
1	0.047	0.015	47,380	0.408	0.065	0.022	17,276	0.512
2	0.056	0.016	43,487	0.414	0.043	0.023	15,316	0.520
3	0.030	0.015	40,324	0.413	0.008	0.022	13,757	0.525
4	0.032	0.018	36,842	0.416	-0.041	0.024	12,267	0.534

D. Structured Product

Horizon	Participants				Nonparticipants			
(s)	β_s	SE	N	$\hat{y}_{\vert L_i=0}$	β_s	SE	N	$\hat{y}_{\vert L_i=0}$
	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
0	0.103	0.016	50,862	0.163	0.130	0.024	19,278	0.040
1	0.105	0.019	47,382	0.172	0.180	0.029	17,276	0.043
2	0.122	0.019	43,489	0.188	0.147	0.029	15,316	0.048
3	0.124	0.018	40,328	0.196	0.097	0.032	13,757	0.051
4	0.099	0.026	36,848	0.206	0.067	0.031	12,267	0.051

Table B.11: Effect of Wealth on Equity Market Participation Probability by Pre-lottery Wealth Quartiles. Coefficients are obtained by estimating Equation 1 in the post-1999 sample stratified by quartiles of pre-lottery net wealth.

	A. Participants			
Quartile:	**1**	**2**	**3**	**4**
Estimate	-0.011	-0.005	0.001	-0.024
SE	0.009	0.016	0.011	0.033
N	478	801	203	50
	B. Nonparticipants			
Quartile:	**1**	**2**	**3**	**4**
Estimate	0.154	0.097	0.114	0.040
SE	0.044	0.035	0.049	0.014
N	9,299	5,387	3,223	1,369

C Online Appendix - Survey Details

C.1 Introduction

This appendix reports descriptive statistics about lottery players' beliefs about stock market returns using a survey that was sent to a subset of the lottery players during the fall of 2016. The survey focused on questions related to well-being, health, and political preferences and beliefs, but also included two questions about stock market returns. The survey was sent to 241 Kombi large-prize winners and 964 (241×4) matched controls, 3,065 Triss-Lumpsum winners and 570 Triss-Monthly winners. To be consistent with the baseline sample analyzed in this paper, we exclude Triss-Monthly winners from all analyzes presented below. Among the Triss-Lumpsum and Kombi lottery players that received the survey, 59 percent at least one of the two questions about stock market returns. Further details about the survey and the exact criteria used when selecting the survey population is provided in Östling, Lindqvist and Cesarini (2016). Data from the same survey has been used in Lindqvist, Östling and Cesarini (2018) and they show there is no evidence that the propensity to answer the survey was related to the amount won.

C.2 Survey Questions and Definitions

The first survey question about stock market returns asks respondents to assess the probability that the market index for the Stockholm Stock Exchange will appreciate during the coming 12-month period. Respondents are given 12 pre-specified response alternatives from 0 to 100 percent. The second question asks respondents to provide an estimate for how much the market index will depreciate or appreciate during the coming 12 months. Respondents could respond with any number betweeen -99 and $+99$ percent. Based on the respondents' answers to these two questions, we define three measures of "extreme" beliefs about the evolution of the stock market.

1. *Extreme subjective probability*. We first define resonable beliefs about the probability of the stock market will appreciate (the first survey question). As a benchmark, we consider the MSCI Sweden index from the previous 20-year period (1996-2016). During these 20 years, the nominal stock market index appreciated 14 years and depreciated 6 years, implying that the stock market rate appreciate 70 percent of all years. We define subjective probabilities outside of 20 percent from this benchmark (i.e., below 50 or above 90 percent) as "extreme".

2. *Extreme expected returns*. We again use the evolution of the MSCI Sweden index 1996-2016 to provide a benchmark for expected returns (the second survey question). The nominal arithmetic average return during this period is 9.6 percent (8.5 percent after adjusting for inflation). We define expected returns below 0 or above +20 percent as "extreme".

3. *Incoherent beliefs*. Our third measure focus on the consistency between the answers to the first and second survey question. In principle, any subjective probability of stock market appreciation between 0 and 100 percent can rationalize any expected return. However, when the subjective probability is close to 0 (100) percent, the distribution of expected return must be very skewed to the right (left) in order for the expected return to be positive (negative). We define beliefs as being "incoherently positive" if the subjective probability is weakly below 25 percent while the expected return is positive and as "incoherently negative" if the subjective probability is weakly above 75 percent while the expected return is negative.

C.3 Descriptive Statistics

We begin by documenting the distributions of beliefs. Because winning the lottery might affect expectations, we restrict attention to non-winners in the Kombi lottery and small-prize winners

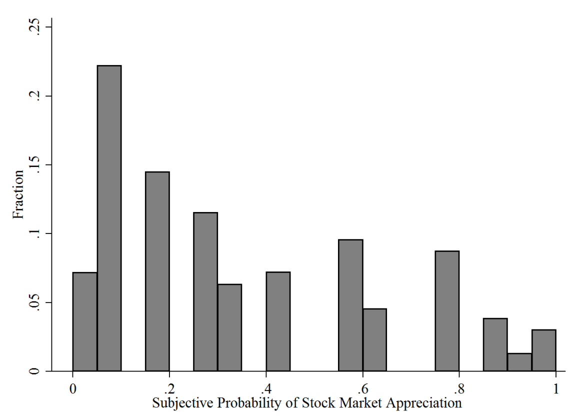
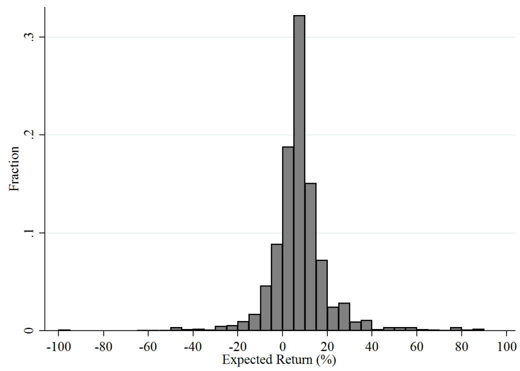


Figure C.1: Histogram of Subjective Probability of Stock Market Appreciation Histogram for the probability the subjects attach to the stock market appreciating during the next 12 months. N = 1,738.

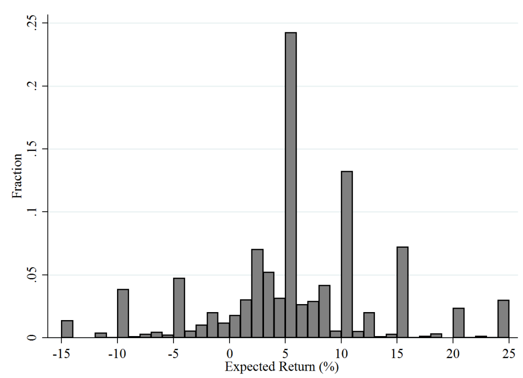
in the Triss lottery (prizes below 150K SEK), reducing the sample to 1,749 individuals. Figure C.1 shows the distribution of subjective probability of stock market appreciation. The distribution is skewed to the right with a median of 25 percent and a mean of 33 percent. According to our definition, 70.4 percent of respondents hold extremely negative beliefs, while only 2.6 percent hold extremely positive beliefs.

Figure C.2(a) shows the distribution of expected returns. To provide a more detailed view, C.2(b) shows the distribution of beliefs when the support of the distribution is limited to returns between -15 and $+25$ percent. Compared to the question about subjective probabilities, the distribution of expected returns is more in line with what well-informed respondents would answer. The average expected return is 5.9 percent, quite close to the close to the historical inflation-adjusted average of 8.5 percent. Only 25.7 percent of respondents hold extreme beliefs about expected returns. As for the subjective probability, unrealistically negative returns are more common (18.8 percent of respondents) than unrealistically positive returns (6.9 percent).

We now turn to incoherent beliefs, i.e., the relationship between expected return and the subjective probabilities. Figure C.3 plots the expected return against the subjective probability with the size of each circle being proportional to the number of respondents. As expected, respondents with a high subjective probability on average report a higher expected return (ρ = 0.29, p < 0.0001), though there is substantial variation in the expected return among respondents who report the same subjective probability. Regressing expected return on the subjective probability, we find an increase in subjective probability by 100 percentage point is associated with a 13.2 percentage



(a) Full Distribution



(b) Limited Support

Figure C.2: Histogram of Expected Returns. Histogram of respondents' expected stock market return for the coming 12 months. N = 1,589 for the full distribution in panel (a).

point increase in the expected return.

Figure C.4 shows the distribution of expected return for respondents whose subjective probability of stock market appreciation is below 25 or above 75 percent, respectively. While the distribution for respondents who attach less than 25 percent chance of the stock market appreciating is clearly shifted to the left, it is noteworthy that the modal expected return is positive also for respondents with low subjective probability. Many respondents who report an unreasonably low subjective probability thus seem to have realistic expectations about stock market returns. A potential explanation for this finding and for the high proportion of respondents who report unreasonably low subjective probabilities is that respondents are used to thinking about the stock market in terms of expected returns and fail to properly grasp what is meant by the probability of the stock market appreciation.

Panel A of Table C.1 shows the different combinations of subjective probability and expected returns. The sample is restricted to the 1,587 individuals who won less than 150K SEK and for whom we observe both the expected return and the subjective probability. The most common combination of beliefs, held by about half of respondents (47.1 percent), is to believe in a realistic expected return but attach an overly negative probability to the stock market appreciating. Only about a quarter of respondents (24.8 percent) hold realistic beliefs about both the expected return and the probability of stock market appreciating. Also common are overly negative views about both the expected return and subjective probability (18.1 percent). Panel B shows the share of respondents who hold incoherent beliefs, as defined in Section II above. While the fraction of respondents

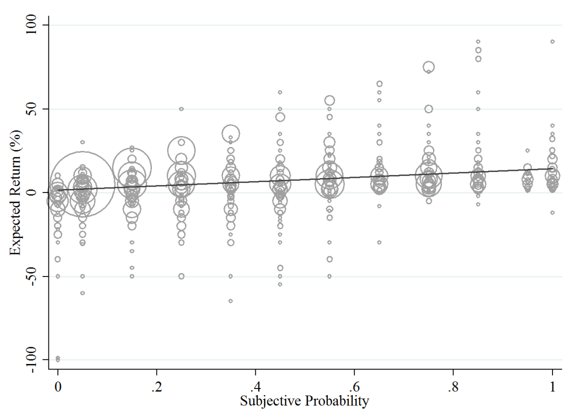


Figure C.3: Expected Return vs. Subjective Probability Each circle indicates a certain combination of subjective probability and expected return. The area of each circle corresponds to the number of respondents within each circle. N = 1,587.

who hold "incoherently negative" beliefs is negligible, more than 40 percent of respondents hold beliefs we classify as "incoherently positive", meaning they state a positive expected return but attach a probability below 25 percent to the possibility of the stock market appreciating. There are two stylized facts to take from the descriptive analyses above. First, irrespective of whether we ask respondents about the expected return or subjective probability that the stock market will appreciate, beliefs are overly pessimistic. Second, a substantial fraction of respondents report answers which are mutually incoherent, suggesting that their basic financial literacy is low. The high fraction of respondents who state a low subjective probability indicate that people may be more used to thinking about the stock market in terms of expected returns.

We now consider which factors predict beliefs about stock market returns. To this end, we regress the different measures of beliefs discussed above on a set of basic demographic characteristics. To test whether beliefs are affected by the stock market return in formative years, we also include the average stock market return between age 18 and 25 as an independent variable. The results from these regressions are reported in Table C.2. We emphasize that our results in this part are purely descriptive and do not have a causal interpretation.

Columns 1-3 of Table C.2 shows the results for the subjective probability. The subjective probability is 9.6 percentage points lower for women, and women are more likely to be overly negative and less likely to be overly positive. The subjective probability is higher for respondents who are born in Sweden, have high labor earnings, many children, and a college degree. The same

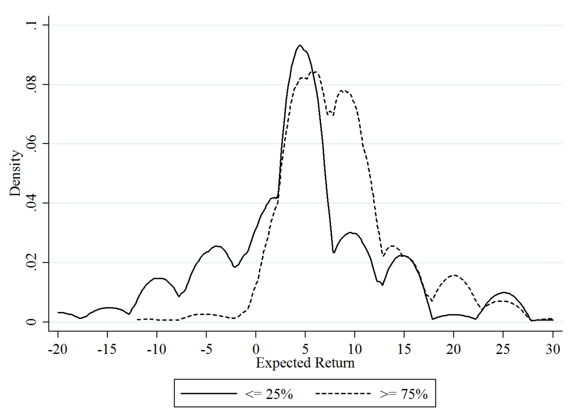


Figure C.4: Expected Return by Subjective Probability The figure shows the Kernel density plots (bandwidth = 1.25) for the distribution of expected returns by the subjective probability that the stock market will appreciate next year. The solid line indicates respondents whose probability is below 25 percent whereas the dashed line indicates respondents with a probability above 75 percent. N = 855 (subjective probability below 25 percent) and N = 271 (subjective probability above 75 percent).

characteristics also predict a lower risk of overly negative beliefs, but are not associated with a higher risk of being overly positive. Columns 4-6 of Table C.2 shows the corresponding results for the expected return. Notably, none of the characteristics that predict the subjective probability predict the expected return. Perhaps surprisingly, women appear more likely to report an overly positive belief about the expected return. Labor income is negatively correlated both with overly positive and overly negative expected returns, indicating that people with high income have a better sense of what returns seem plausible. The stock market return in formative age appear to be unrelated to both the subjective probability and the expected return.

While we saw above that respondents, in general, are too pessimistic about stock market returns, no characteristics except labor income predict overly negative beliefs for both the subjective probability and the expected returns. One explanation is that pessimistic beliefs are simply weakly correlated with basic socio-economic characteristics. Yet, if so, it remains to explain why the subjective probability is quite strongly related to several socio-economic characteristics, so one possibility is that many respondents have trouble understanding the question about subjective probability, and that the correlations between the subjective probability and socio-economic characteristics are due to low financial literacy rather than pessimism per se. Consequently, the large fraction of respondents who report overly negative beliefs about the subjective probability in Table C.1 may thus at least partly reflect a lack of financial literacy or basic numeracy skills.

Table C.1: Frequency of Beliefs. Sample of non-winners and winners of prizes below 150K SEK for which both subjective probability and expected return is observed. N = 1,587.

Panel A. Combination of Beliefs

Percent of respondents	Expected returns			
	Overly positive (>+20%)	Realistic (0% to +20%)	Overly negative (<0%)	Sum
Subjective probability				
Overly positive (>90%)	0.3	2.4	0.1	2.8
Realistic (50% to 90%)	2.7	24.8	0.6	28.2
Overly negative (<50%)	3.9	47.1	18.1	69.1
Sum	6.9	74.3	18.8	100

Panel B. Coherent vs Incoherent Beliefs

Percent of respondents	Coherent	Incoherently positive	Incoherently negative	Sum
	59.2	40.5	0.3	100

This interpretation is broadly consistent with previous literature which has revealed lower financial literacy among women (Lusardi and Mitchell 2008; Almenberg and Dreber 2015) and for people with low education (van Rooij, Lusardi and Alessie 2011). The results in Column 7 of Table C.2 – where we regress an indicator variable for incoherent beliefs on the same set of socioeconomic characteristics – lend some support to this view. Female respondents are 12.5 percentage points more likely and respondents with a college degree 15.6 percentage points more likely to report incoherent beliefs, and both effects are strongly statistically significant.

C.4 Conclusion

Our analysis of the survey responses show that for both our survey questions aimed at eliciting the respondents' beliefs, a large fraction of respondents reports pessimistic beliefs about the evolution of the stock market. Moreover, both the large number of respondents who attach a probability close to zero to the stock market appreciating, and the large fraction whose responses are appear mutually incoherent, suggest financial literacy is low in our sample. The propensity to state extreme or incoherent beliefs also correlate with socio-economic factors which previous literature has shown predict low financial literacy.

Table C.2: Predictors of Stock Market Return Beliefs. All regressions estimated with OLS. The sample is restricted to non-winners and winners of 150K SEK or less. All time-varying independent variables refer to year 2014. Labor and Capital income have been winsorized at the 1st and 99th percentile. Robust standard errors are reported in parenthesis and the corresponding *p*-values in brackets.

	Subjective probability			Expected return			
	Stated probability (0-1)	Overly positive (0/1)	Overly negative (0/1)	Stated return (%)	Overly positive (0/1)	Overly negative (0/1)	Incoherent beliefs (0/1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.096	-0.021	0.147	0.115	0.027	0.023	0.125
	(0.014)	(0.008)	(0.022)	(0.724)	(0.014)	(0.020)	(0.025)
	[0.000]	[0.009]	[0.000]	[0.874]	[0.049]	[0.251]	[0.000]
Age/10	-0.091	-0.010	0.122	-1.710	0.029	0.095	0.028
	(0.046)	(0.028)	(0.072)	(2.256)	(0.042)	(0.066)	(0.081)
	[0.048]	[0.715]	[0.089]	[0.449]	[0.490]	[0.153]	[0.733]
Age/10	0.007	0.000	-0.010	0.231	-0.003	-0.013	0.002
squared	(0.005)	(0.003)	(0.008)	(0.245)	(0.005)	(0.007)	(0.009)
	[0.160]	[0.973]	[0.186]	[0.346]	[0.476]	[0.082]	[0.817]
Married	0.004	-0.001	0.011	0.035	-0.001	-0.011	0.025
	(0.014)	(0.008)	(0.022)	(0.667)	(0.013)	(0.020)	(0.025)
	[0.747]	[0.869]	[0.603]	[0.959]	[0.928]	[0.600]	[0.314]
Born in	0.063	0.011	-0.100	0.962	-0.001	-0.002	-0.058
Sweden	(0.027)	(0.013)	(0.039)	(1.464)	(0.026)	(0.041)	(0.050)
	[0.019]	[0.384]	[0.010]	[0.511]	[0.982]	[0.956]	[0.251]
Capital income	0.019	0.004	-0.032	0.375	-0.009	-0.015	-0.011
(in 100K SEK)	(0.014)	(0.006)	(0.022)	(0.614)	(0.008)	(0.016)	(0.023)
	[0.184]	[0.524]	[0.144]	[0.541]	[0.285]	[0.335]	[0.645]
Labor income	0.021	-0.001	-0.029	0.058	-0.011	-0.022	-0.011
(in 100K SEK)	(0.005)	(0.003)	(0.008)	(0.245)	(0.004)	(0.007)	(0.008)
	[0.000]	[0.799]	[0.000]	[0.813]	[0.009]	[0.001]	[0.200]
Number of	0.020	-0.002	-0.035	0.802	0.007	-0.020	-0.020
children	(0.008)	(0.005)	(0.014)	(0.419)	(0.008)	(0.012)	(0.015)
	[0.018]	[0.613]	[0.010]	[0.056]	[0.416]	[0.103]	[0.180]
College degree	0.085	0.006	-0.136	-0.846	-0.015	0.006	-0.156
	(0.016)	(0.009)	(0.026)	(0.788)	(0.014)	(0.024)	(0.027)
	[0.000]	[0.502]	[0.000]	[0.283]	[0.271]	[0.800]	[0.000]
Stock market return	-0.035	-0.009	-0.001	-5.363	-0.054	0.197	-0.269
in formative age	(0.124)	(0.072)	(0.194)	(5.730)	(0.113)	(0.179)	(0.215)
	[0.779]	[0.901]	[0.997]	[0.349]	[0.635]	[0.273]	[0.210]
R^2	0.092	0.009	0.080	0.005	0.012	0.016	0.066
N	1,736	1,736	1,736	1,596	1,596	1,596	1,585