WINDFALL GAINS AND STOCK MARKET PARTICIPATION

Joseph S. Briggs
David Cesarini
Erik Lindqvist
Robert Östling

Windfall Gains and Stock Market Participation
Joseph S. Briggs, David Cesarini, Erik Lindqvist, and Robert Östling
NBER Working Paper No. 21673
October 2015
JEL No. D1,G02,G11

ABSTRACT

We estimate the causal effect of wealth on stock market participation using administrative data on Swedish lottery players. A $150,000 windfall gain increases stock ownership probability among pre-lottery non-participants by 12 percentage points, while pre-lottery stock holders are unaffected. The effect is immediate, seemingly permanent and heterogeneous in intuitive ways. Standard lifecycle models predict wealth effects far too large to match our causal estimates under common calibrations. Additional analyses suggest a limited role for explanations such as procrastination or real-estate investment. Overall, results suggest that "nonstandard" beliefs or preferences contribute to the nonparticipation of households across many demographic groups.

Joseph S. Briggs
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10012
jsb493@nyu.edu

David Cesarini
Department of Economics
New York University
19 West 4th Street, Room 511
New York, NY 10012
and NBER
dac12@nyu.edu

Erik Lindqvist
Stockholm School of Economics
Box 6501
SE-113 83 Stockholm
SWEDEN
erik.lindqvist@hhs.se

Robert Östling
Stockholm University
Institute for International Economic Studies
106 91 Stockholm
Sweden
robert.ostling@iies.su.se

Windfall Gains and Stock Market Participation[*]

By Joseph Briggs,[†] David Cesarini, Erik Lindqvist, and Robert Östling

We estimate the causal effect of wealth on stock market participation using administrative data on Swedish lottery players. A $150,000 windfall gain increases stock ownership probability among pre-lottery non-participants by 12 percentage points, while pre-lottery stock holders are unaffected. The effect is immediate, seemingly permanent and heterogeneous in intuitive ways. Standard lifecycle models predict wealth effects far too large to match our causal estimates under common calibrations. Additional analyses suggest a limited role for explanations such as procrastination or real-estate investment. Overall, results suggest that "nonstandard" beliefs or preferences contribute to the nonparticipation of households across many demographic groups.

"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." - Carroll (2002)

Canonical life-cycle models of consumption and savings predict all individuals should invest a positive fraction of their wealth in equities (Samuelson (1969); Merton (1971)). However, a sizeable 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 what has been dubbed the "non-participation puzzle" (Haliassos and Bertaut (1995); Vissing-Jørgensen (2003); Campbell (2006); Guiso and Sodini (2013)). Insights into the causes of non-participation may guide efforts to more effectively promote efficient financial decision making (Campbell (2006)).

Limited stock market participation is often analyzed in models where agents weigh the benefits of participation against its costs (Mulligan and Sala-i Martin (2000); Vissing-Jørgensen (2002); Vissing-Jørgensen (2003); Paiella (2007); Attanasio and Paiella (2011)). A pioneer in this literature was Vissing-Jørgensen (2003), who proposed 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 and plausible structural interpretation of the robustly documented positive correlation between wealth and stock market participation (Mankiw and Zeldes (1991); Poterba and Samwick (2003); Campbell (2006)). Vissing-Jørgensen (2003) showed that in a model calibrated using the US cross-sectional wealth distribution, per-period participation costs of a magnitude comparable to realistic estimates of the direct financial costs of participation could account for the majority of the non-participation of all but the wealthiest households.

This framework has proven to be a valuable foundation for subsequent work, much of which extends and refines the theory along several dimensions. Housing (Cocco (2005); Flavin and Yamashita (2011); Vestman (2013)) and outstanding debt (Davis, Kubler and Willen (2006); Becker and Shabani (2010)) reduce benefits of participation by providing a potentially preferred investment. Private business equity (Heaton and Lucas (2000a)) and stochastic labor income (Viceira (2001)) increase background risk and may be positively correlated with equity returns, thus reducing benefits of participation.

Though many portfolio theories make precise, quantitative predictions about the effect of a windfall gain on risk-taking behavior, credibly testing these predictions is difficult, as the opening quote by Carroll illustrates. In this paper, we estimate the causal effect of wealth on stock market participation 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 (e.g., number of tickets owned) conditional on which the lottery wealth is randomly assigned. Second, because the size of the prize pool is over 500 million dollars, 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 analyses proceed in two stages. We begin by reporting the results from a set of reduced-form analyses of how wealth affects participation. On average, a positive net wealth shock of 1M SEK (approximately 150K USD) increases the participation probability in post-lottery years by 4 percentage points.[1] This effect is accounted for entirely by a 12 percentage-point increase in the stock market participation of households that did not participate in equity markets at the time of the lottery. The positive effect in these households is precisely estimated, immediate, seemingly permanent, and heterogeneous in directions that are easy to reconcile qualitatively with the predic-

---

[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/1 USD.

tions of standard models: wealth effects are larger in households who are poorer, college-educated, debt-free, not self-employed, and win during a period of positive returns. We also document that in nonparticipants, the effect of wealth on bond-market entry is greater than its effect on equity entry, both on the intensive and extensive margin.

In the second stage, we examine how our results help narrow the set of hypotheses about the causes of non-participation. We show in a model validation exercise that under a wide range of calibrations, a standard life-cycle model systematically predicts effects of windfall gains on participation substantially larger than those we estimate. We also estimate the fixed entry and participation costs needed to match our causal estimates. Our primary conclusion is that over 45 percent of nonparticipants are estimated to have entry costs greater than 1M SEK (150K USD), whereas per-period participation costs are quite modest and exhibit little variation across households. When we introduce heterogeneity by pre-lottery characteristics, we find that in all subpopulations considered, the median entry costs are much larger than typical estimates in the structural literature (e.g., Gomes and Michaelides (2005); Cocco (2005); Alan (2006); Khorunzhina (2013)).

A first contribution of this paper is to provide credible and precise estimates of how large wealth shocks affect stock market participation. Researchers interested in the relationship between wealth and participation are usually confined to studying observational data (Calvet, Campbell and Sodini (2007); 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. Our research design allows us to closely approximate 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. This approach is most similar to a quasi-experimental study by Andersen and Nielsen (2011), which uses Danish data on inheritances due to sudden parental deaths as a source of wealth variation.[2] Despite differences in sample and research design, our reduced-form estimates are broadly similar to theirs.

A second contribution is our use of a structural model to estimate the costs needed to rationalize observed participation responses. This exercise facilitates the quantitative comparison of our quasi-experimental estimates to a comprehensive body of structural work (e.g., Gomes and Michaelides (2005); Cocco (2005); Alan (2006); Khorunzhina (2013); Fagereng, Gottlieb and Guiso (2015)). Looking at patterns and magnitudes of the estimated costs may help refine theories of participation, evaluate the quantitative importance of various mechanisms that have been proposed, and provide cues about promising avenues for future modeling efforts.

Non-participation of the wealthiest households has been previously described as a significant

---

[2]Lotteries provide a source of variation that is unlikely to be correlated with other significant life events that could influence financial choices. In contrast, sudden parental death generates an initial wealth shock, but also eliminates any future intergenerational transfers, including future inheritance.

challenge to financial theory" (Campbell (2006), p. 1564) because, under standard calibrations, most models imply that households forgo large welfare gains by declining to own stocks. Our results suggest that the challenge extends to a substantial fraction of non-wealthy households too. Even windfall gains in excess of 2M SEK (300K USD) do not cause a majority of nonparticipating households to enter equity markets, and an unwillingness to enter is pervasive across all considered subpopulations. Accounting for this non-participation requires costs much larger than are typically considered in the structural literature. Additional analyses suggest that many previously hypothesized extensions, such as allowing for uninsurable income risk, housing, and procrastination, are unlikely to fully explain our results. Overall, our results suggest that cognitive constraints or "non-standard" beliefs and preferences are likely to play an important role in explaining the behavior of nonparticipating households (Vissing-Jørgensen (2003); Ang, Bekaert and Liu (2005); Barberis, Huang and Thaler (2006); Guiso, Sapienza and Zingales (2008); Biais, Bossaerts and Spatt (2010); Campanale (2011); Grinblatt, Keloharju and Linnainmaa (2011)).

The remainder of the paper is structured as follows. Section 1 describes the lottery and wealth data, our identification strategy, and addresses several issues regarding external validity that are often raised about studies of lottery players. In Section 2, we report the results from our reduced-form analyses. Section 3 presents a structural model of life-cycle asset market participation, examines its ability to account for our causal estimates, and estimates the distribution of entry and participation costs implied by our variation in wealth. Finally, Section 4 discusses our findings in the context of the literature on nonparticipation and concludes.

1   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).

1.1   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 aggregate assets and debt, and relevant subcategories such as bank account balances, mutual funds, directly held stocks, bonds, money market funds, debt, residential and commercial real estate, and other financial and real assets. The data have proven valuable in household-finance research beginning with a landmark paper by Calvet et al. (2007). The quality of the Wealth Register compares favorably to other data sources, as it is not subject to attrition and included variables are likely measured with minimal error. Calvet et al. (2007) estimate that included variables account for approximately 86% of wealth in Sweden.

We supplement the portfolio data from the Wealth Register with basic demographic information available in the Statistics Sweden administered database LISA. Our analyses are conducted at the household level, with a household defined as the observed winner and, if present, his or her spouse. We choose this definition because the wealth of spouses of winning players increases by about 10% of the prize won following the lottery event, thus suggesting some joint control over assets. All our analyses are based on players aged 18 and above, and we restrict the sample to lottery draws conducted no later than 2007, the last year for which we have financial data. In addition, given that we consider a maximum horizon of 10 years and wealth data is only available starting in 1999, we restrict our sample to prize winners after 1989.

1.2 Lottery Data

Our basic 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, we describe each lottery separately. For a detailed description of how the original lottery data were preprocessed and quality-controlled, we refer the reader to the Online Appendix of Cesarini, Lindqvist, Östling and Wallace (2015).

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. Subscribers choose their desired number of subscription tickets and are billed monthly, usually by direct debit. 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.

The Kombi rules are simple. In a given 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 349 large prize-winners, matched to a total of 34,852 controls.

Triss –   The second sample is a scratch-ticket lottery run since 1986 by Svenska Spel, the Swedish government-owned gambling company. Since 1994, Triss lottery players can win the opportunity to participate in a TV show where they can win substantial prizes. In a typical month, 25 Triss winners appear on the show and draw a prize by selecting a ticket from a stack. The tickets are and shuffled and look identical. The prizes are distributed according to a known prize plan with prizes

varying from 50K SEK to 5M SEK. The prize plan is subject to occasional revision.

Svenska Spel supplied the basic demographic information (name, age, region of residence, and often also the names of close relatives) about all individuals who participated in the TV show between 1994 and 2010. With the help of Statistics Sweden, we were able to reliably identify the PINs of 99% of show participants. Svenska Spel also listed cases in which the player shared ownership of the ticket. Our analyses are based exclusively on the 90% of winners who did not indicate they shared ownership of the winning ticket. Our empirical strategy makes use of the fact that, conditional on the prize plan, the nominal prize amount is plausibly random. Thus, two players are assigned to the same cell if they won in the same year and under the same prize plan, providing a final sample of 3,416 winners.

PLS – 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. Such accounts have existed in Sweden since 1949 and were originally subsidized by the government. When the subsidies ceased in 1985, the government authorized banks to continue to offer prize-linked-savings products. Two systems were put into place, one operated by savings banks and one by all other banks. The two systems were approximately equally popular and participation was widespread across broad strata of Swedish society; every other Swede owned an account.

The PLS sample was obtained by combining data from two sources of information about the PLS accounts maintained by the commercial banks and state bank. The first source is a set of printed lists with information about prizes won in the draws between 1986-2003. For each prize won in a draw, these sheets list the prize amount, type of prize won (described below), and the winning account number. The second source is a large number of microfiche images with information (account number, account owner's PIN, and number of tickets received) about all eligible accounts participating in the draws between December 1986 and December 1994 (the "fiche period"). Because the prize lists contain the winning account number, but not its owner PIN, the fiches are needed to identify winning players' PIN.

PLS account holders could win two types of prizes: odds prizes and fixed prizes. The probability of winning either type of prize was proportional to the number of tickets associated with an account: account holders assigned one lottery ticket per 100 SEK in account balance. Fixed prizes, which constitute the majority of 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 sum of fixed prizes won is independent of the account balance. Previous studies of lottery players have used this identification strategy (Imbens, Rubin and Sacerdote (2001); Hankins, Hoestra and Skiba

Table 1: Overview of Identification Strategy.

Lottery	Period	Prize Type	Cells
PLS	1990-2003	Fixed Prize	Draw × # Fixed Prizes
PLS	1990-1994	Odds Prize	Draw × Balance
Kombi	1994-2007	Fixed Prize	Draw × # Tickets
Triss	1994-2007	Fixed Prize	Year × Prize Plan

(2011)). Because the strategy does not require information about the number of tickets owned, it can be employed also during the post-fiche period, as long as the winning account was active during the fiche period so the account owner's PIN can be identified. We therefore assign two individuals to the same cell if they won an identical number of fixed prizes in that draw. Overall, we were able to reliably match 99% of the fixed-prize-winning accounts from the fiche era to a PIN.

To construct odds-prize cells, we match individuals who won exactly one odds-prize 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. After the fiche period, we do not observe account balances and therefore odds prizes are only included if won during the fiche period (1986-1994). In total, the sample includes 275,248 PLS prizes, of which 257 are larger than 1M SEK.

1.3 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,

$$(1) \qquad Y_{i,s} = \beta_s \times L_{i,0} + \mathbf{X}_i \times \mathbf{M_s} + \mathbf{Z}_{i,-1} \times \boldsymbol{\gamma_s} + \eta_{i,s},$$

where i indexes households, $L_{i,0}$ denotes the prize size (in million SEK), $\mathbf{X}_i$ is a vector of cell fixed effects, and $\mathbf{Z}_{i,-1}$ is a vector of controls. Controls are included to improve the precision of our estimates and are always measured in the year before the lottery. 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.

In many of our analyses, we estimate Equation 1 in our pooled sample. We also conduct several analyses 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

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

Prize Amount	A. All-Year				B. Post-1999			
	Pooled	PLS	Kombi	Triss	Pooled	PLS	Kombi	Triss
$L_i \leq 10K$	293,470	258,618	34,852	0	71,211	42,223	28,988	0
$10K < L_i \leq 100K$	16,020	15,031	0	989	742	375	0	367
$100K < L_i \leq 500K$	3,348	1,267	0	2,081	1,240	0	0	1,240
$500K < L_i \leq 1M$	232	75	0	157	89	0	0	89
$1M < L_i \leq 2M$	605	222	333	50	298	2	274	22
$2M < L_i$	190	35	16	139	78	0	16	62

regressions, we control for the following lagged baseline demographic characteristics: age, sex, marital status, retirement status, college completion, household size, disposable income, and self-employment. In the post-1999 sample regressions, we additionally control for the following lagged baseline financial characteristics: net wealth, gross debt, and indicators for real estate ownership and participation (directly held stocks or mutual funds).

Prize Variation – To get a better sense of the source of our identifying variation, Table 2 provides information about the distribution of prizes. The total value of the after-tax prize money disbursed to the winners in our samples is almost 3.5 billion SEK (about 500M USD), 67% of which is accounted for by prizes whose value is greater than the median annual Swedish household disposable income in 1999 (143K SEK (21K USD)). Thus, although small prizes account for a relatively large fraction of prizes won, most identifying variation comes from the larger prizes in all three lotteries. All lotteries contribute substantial identifying variation to the all-year sample, whereas Kombi and Triss prizes jointly account for most identifying variation to the post-1999 sample.

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:

$$(2) \qquad L_{i,0} = \mathbf{X_i} \times \mathbf{\Gamma_0} + \mathbf{Z_{i,-1}} \times \boldsymbol{\rho_{-1}} + \epsilon_i.$$

Under the null hypothesis of conditional random assignment, the characteristics determined before the lottery ($\mathbf{Z_{i,-1}}$) should not predict the lottery outcome (L_i) conditional on the cell fixed ($\mathbf{X_i}$) effects. We run these quasi-randomization tests in the all-year sample, its lottery subsamples, and 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 7), the null hypotheses of random assignment is rejected.

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

| | **All-Year** | | | | | **Post-1999** | |
| | **Pooled** | | **PLS** | **Kombi** | **Triss** | **Pooled** | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fixed Effects	Cells	None	Cells	Cells	Cells	Cells	None
Demographic Controls							
F-stat	.69	11.54	.64	.92	1.37	.87	10.01
p	.74	<.001	.78	.51	.19	.56	<.001
Financial Controls							
F-stat	—	—	—	—	—	1.81	12.80
p	—	—	—	—	—	.14	<.001
Demographic+Financial Controls							
F-stat	—	—	—	—	—	1.29	15.20
p	—	—	—	—	—	.22	<.001
N	313,865	313,865	275,248	35,201	3,416	73,658	73,658

1.4 Generalizability

In this section we address two important concerns about the external validity of our sample. A first concern is that individuals who play the lottery may not be representative of the population at large. A second is that inferences from Swedish lottery players about the causes of non-participation may not generalize to other countries.

Generalizing within Sweden – To investigate the representativeness of our samples, 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 the sex- and age- matched representative sample. In the matched population sample, 63% of households are home owners and 56% participate in equity markets; in our pooled sample, the corresponding figures are 70% and 59%. Columns 5-7 provide the corresponding descriptive statistics for the post-1999 sample broken down by lottery.

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

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-Year | | Post-1999 | | | | |
| | Pooled | Pop | Pooled | Pop | PLS | Kombi | Triss |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Demographic							
Female	.504	.504	.516	.516	.575	.436	.558
Age (years)	56.7	56.7	56.3	56.3	63.2	62.2	51.9
Household Members (#)	1.92	1.94	1.97	1.97	1.75	1.81	2.13
Household Income (K SEK)	212	208	253	250	189	210	288
Married	.542	.552	.519	.525	.518	.483	.543
Retired	.353	.342	.311	.279	.481	.425	.217
Self-Employed	.050	.048	.046	.059	.026	.003	.040
Student	.033	.033	.026	.032	.032	.078	.052
College	.179	.204	.193	.257	.229	.153	.216
Financial							
Net Wealth (K SEK)			879	1085	1,480	830	851
Gross Debt (K SEK)			362	347	235	246	452
Home Owner			.702	.630	.666	.732	.686
Equity Participant			.591	.558	.682	.625	.560
N	313,442	49,959	73,644	49,959	42,586	29,278	1,780

cross-sectional probit equation used by Calvet et al. (2007) in their study of a large sample of representative Swedes. To avoid including wealth variation that was induced by the lottery, we restrict the estimation sample to the post-1999 sample and estimate the probit specification using the 1999 cross-sectional wealth data in (i) our post-1999 lottery sample and (ii) a sex- and age-weighted representative sample. The results, reported in Appendix Table C.2, show the overall pattern of conditional correlations are similar in our lottery sample and the reweighted representative sample.

Generalizing beyond Sweden –   The processes that cause participation in Sweden may differ in important ways from the processes in other countries. An indirect way to evaluate generalizability beyond Sweden is to compare the cross-sectional relationships of financial variables with demographic characteristics in Sweden to other countries. Previous work has noted that the predictors of non-participation in Sweden and the United States are similar, as is the aggregate composition of household wealth in the two countries. For example, the Swedish participation rate was 62% in 1999, compared to 59% in the United States (Campbell (2006), pp. 1572-1576).

   To provide further indirect evidence on generalizability we compare the distribution of participation costs implied by the 1999 cross-sectional wealth data in our post-1999 lottery sample with
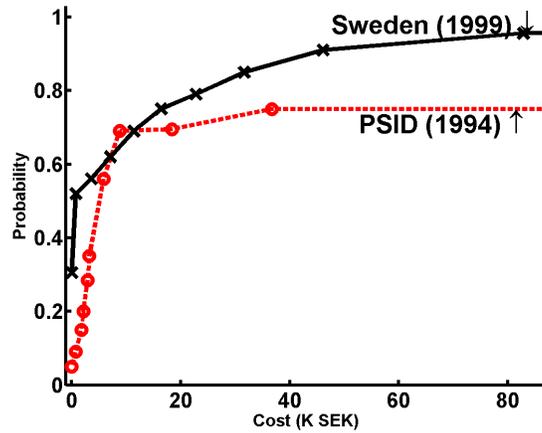
Figure 1: Comparing CDFs of Estimated Per-Period Participation Costs in Sweden and the US. For methodological details, see Exercise B of Vissing-Jørgensen (2003).

the distribution calculated by Vissing-Jørgensen (2003) using the 1994 PSID. In Vissing-Jørgensen (2003)'s framework, households decline to participate if at their level of wealth, the gains from participation are too small to offset the costs. The benefits are the expected equity premium for the share of the wealth that the household chooses to allocate to the risky financial portfolio. Assuming time separable and homothetic preferences, the per-period benefit of participation of household i at time t can be approximated by

$$(3) \qquad Benefit_{i,t} = W_{i,t} \times \alpha_{i,t} \times (r_{i,t}^{ce} - r_t^f)$$

where $W_{i,t}$ is household i's wealth at time t, $\alpha_{i,t}$ is the fraction of wealth household i would invest in equities at time t if they participate, and $r_{i,t}^{ce}$ is the certain return that would make a household indifferent between investing in risky equity and investing in an asset commanding a certain return of $r_{i,t}^{ce}$. An estimate of a households participation cost is then obtained from the dollar amount required to offset this benefit. Figure 1 shows that applying Vissing-Jørgensen (2003)'s methodology with $(r_{i,t}^{ce} - r_t^f) = .04$ results in similar CDFs of participation costs in the two samples.

2 Empirical Results

We now turn our attention to our reduced form analyses. Our primary outcome variable is year-end participation, defined as an indicator equal to 1 for individuals who own stocks, either directly or indirectly. Figure 2 presents the estimated coefficients for $s = -1, ..., 10$ from the pooled lottery sample. We estimate that 1M SEK causes a near-immediate and permanent increase in the participation probability of around 4 percentage points. As expected, lottery wealth does not

predict participation prior to the lottery. The results are similar if we define participation more narrowly to only include directly owned stocks (Appendix Table C.3, Panel B).
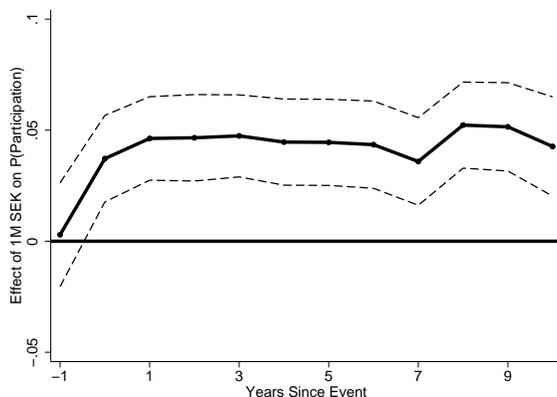
Figure 2: Effect of Wealth (1M SEK) on Participation Probability. Coefficients and 95% confidence intervals, are obtained by estimating Equation 1 in the all-year sample. See Appendix Table C.3 for the underlying estimates.

We next investigate if the effects estimated in the pooled sample mask any treatment-effect heterogeneity. An obvious potential source of heterogeneity is the pre-lottery participation status. Figure 3 shows the estimated treatment-effects on participation probability in $s = -1, ..., 4$ in the post-1999 lottery sample stratified by pre-lottery participation status. The estimated effect of wealth differs dramatically by pre-lottery participation status. In pre-lottery nonparticipants, we estimate that 1M SEK increases the participation probability by 12.1 percentage points. In pre-lottery participants, the estimated effect is small and usually not statistically distinguishable from zero. Hence, the aggregate effect of 3.7 percentage points we observe in the pooled sample appears to be driven nearly entirely by a positive effect on nonparticipants. This finding is consistent with the predictions of a model in which large, one-time, fixed costs of entry are a major cause of non-participation (where "cost" is interpreted broadly to include not just financial costs). 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.[3]

Given that the causal effects appear to be driven entirely by positive effects on pre-lottery nonparticipants, we conducted a suite of additional heterogeneity analyses in this subsample, stratifying the nonparticipants by pre-lottery debt, home ownership, net wealth, stock returns in the calendar year of win, self-employment, sex, age, and educational attainment.[4] Results from these

---

[3]There are two reasons for the widening of the confidence intervals. First, because participation is only observed during a nine year period and we condition on prior participation status, the sample size decreases as we expand the time horizon. Second, the predictive power of the lagged financial and demographic characteristics fall with time, increasing the standard errors.

[4]Procedurally, we run a single regression in which all regressors are interacted with an indicator variable for one

(a) Nonparticipants   **(b) Participants**
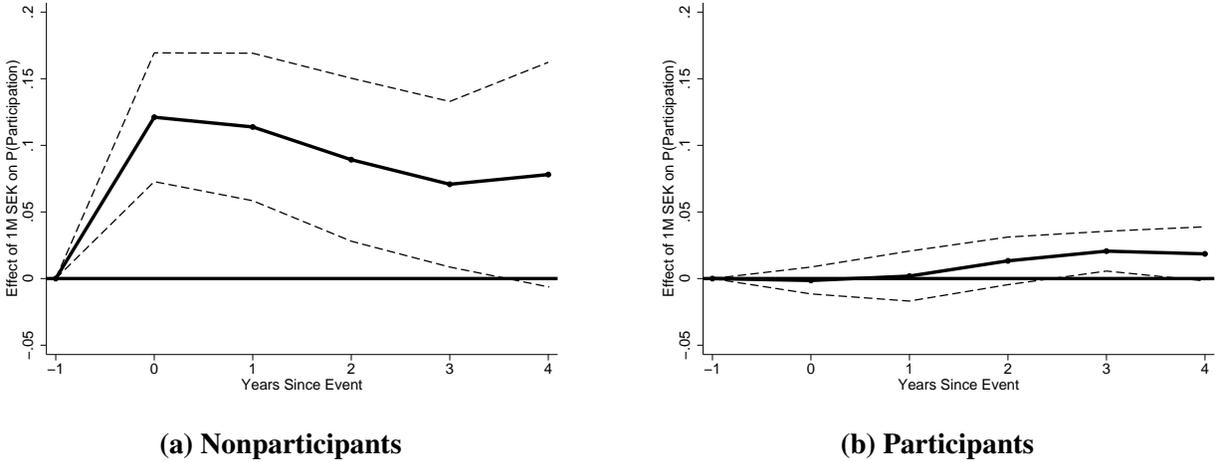
Figure 3: Effect of Wealth (1M SEK) on Participation Probability by $s = -1$ Participation Status.
Coefficients and 95% confidence intervals, are obtained by estimating Equation 1 in the post-1999 sample.
See Appendix Table C.4 for the underlying estimates.

heterogeneity analyses for $s = 0$ are presented in Table 5.[5]

The heterogeneity analyses provide information about how participation costs are distributed across households with different observable characteristics, but are subject to the important interpretational caveat that only wealth is randomly assigned. To illustrate, consider the heterogeneity analyses by pre-lottery debt which shows the estimated effect on participation probability is about twice as large in debt-free households. One theoretical mechanism that could account for this finding is that indebted households face interest rates that substantially increase the interest-free rate (Davis et al. (2006); Becker and Shabani (2010)), making debt reduction a more attractive way to spend the windfall gain than stock market entry. This is a plausible explanation of the observed heterogeneity, but our data do not allow us to rule out the possibility that the observed heterogeneity arises because debt is correlated with other factors that shift participation incentives.

Next, we examine whether the effects differ by pre-lottery ownership of real estate. Theoretically, the net effect of real estate ownership on participation incentives is ambiguous. For example, homeowners may find participation less attractive because they have access to investment opportunities in their home, or prefer to pay off mortgage debt with higher interest rates. On the other hand, as several studies highlight (Grossman and Laroque (1990); Cocco (2005); Flavin and Yamashita (2011); Vestman (2013)), non-real estate owners may use the windfall gain to invest in real estate and this could crowd out stock purchases. In practice, we estimate effects of similar magnitude in owners and non-owners of real estate. When we examine if the effects are heterogeneous by

---

of the subpopulations. The resulting coefficient estimates are identical to those obtained when Equation 1 is estimated separately in each subsample.

[5]Results for $s = 3$ are shown in Appendix Table C.5 and are broadly similar but less precisely estimated because the estimation sample size shrinks with time horizon.

Table 5: Heterogeneous Effect of Wealth (1M SEK) on Participation Probability. Coefficients are obtained by estimating Equation 1 in the post-1999 sample of nonparticipants at time $s = 0$. 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.

	Financial Characteristics							
	Gross Debt		**Home Owner**		**Net Wealth**		**Equity Returns**	
	≤ 0	**> 0**	**Yes**	**No**	**High**	**Low**	**≤ 0**	**> 0**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Effect	.212	.094	.105	.146	.062	.138	.056	.139
SE	(.036)	(.026)	(.027)	(.050)	(.035)	(.029)	(.039)	(.029)
p	<.001	<.001	<.001	.004	.081	<.001	.152	<.001
Hetero p	.008		.476		.094		.087	
N	10,059	10,229	10,649	9,639	4,876	15,412	10,841	9,447

	Demographic Characteristics							
	Sex		**Age**		**College**		**Self-Employed**	
	Male	**Female**	**≤ 45**	**> 45**	**Yes**	**No**	**Yes**	**No**
	(9)	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
Effect	.149	.117	.141	.113	.232	.097	.023	.134
SE	(.038)	(.031)	(.042)	(.031)	(.051)	(.026)	(.024)	(.026)
p	<.001	<.001	.001	<.001	<.001	<.001	.348	<.001
Hetero p	.521		.599		.019		.002	
N	9,251	11,037	2,724	17,564	5,095	15,193	679	19,609

pre-lottery wealth, we find that households with below-median financial wealth are more affected.

Fourth, we ask if the effect differs in years with positive equity returns, as indeed it may if individuals overweight recent events when forming subjective beliefs about future returns. Survey research has found investors and chief financial officers adjust their beliefs about the one-year equity premium downward (upward) following periods of negative (positive) market returns (Graham and Harvey (2001); Vissing-Jørgensen (2003); Greenwood and Shleifer (2014)). We estimate that in households who win the year after a year with negative equity return, 1M SEK increases the participation probability by .056, compared to .139 for households who win the year after a year with positive returns.

Our last four dimensions of heterogeneity are self-employment, age, sex, and educational attainment. If the self-employed face greater uninsurable idiosyncratic labor income risk than do regular salaried employees, the standard life-cycle model predicts the self-employed benefit less from participation (Heaton and Lucas (2000b); Viceira (2001)). Consistent with this hypothesis, we find no evidence of a positive effect of wealth on the participation probabilities of the self-employed. In our age analyses, our estimated effects are larger in younger players, but the

difference is not statistically significant. We similarly do not find a statistically significant difference in the effect on the participation probabilities of men and women, but do find significant differences by education level. The treatment-effect for college-educated households is twice as large as for households without a college degree. One plausible interpretation of this finding is that households with college degrees face smaller information costs or less psychological discomfort (Grinblatt et al. (2011); Van Rooij, Lusardi and Alessie (2012); Benjamin, Brown and Shapiro (2013)). Alternatively, Calvet et al. (2007) show that educated households diversify their portfolios more effectively and hence gain more from entry.

Overall, the results from these heterogeneity analyses show that the effect of wealth on stock market participation generally varies in intuitive ways that are easy to reconcile qualitatively with many of the proposed theories of non-participation.

2.1 Are the Effects Nonlinear?

Under our identifying assumption, our linear estimator gives an unbiased estimate of a weighted treatment-effect. Because large prizes account for most of our identifying variation, the estimator will assign most weight to the marginal effect of wealth at modest to large levels of wealth. To test for nonlinear effects, we modify our basic estimating equation, replacing the lottery-wealth variable by indicator variables for prizes in five categories. We then run regressions in which the smallest prize category omitted.

Figure 4 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 (<10K SEK, 1.5K USD), a prize in the range 10K-100K SEK (1.5 to 15K USD) increases the participation probability of pre-lottery nonparticipants by .013. The corresponding estimates for winners of prizes in the 100K-1M (15 to 150K), 1M-2M (150 to 300K), and 2M+ (300K+) are 0.080, 0.183 and 0.392. Thus, the marginal effect is everywhere positive, but strongest for winners of small prizes. Among pre-lottery participants, none of the prize-category coefficients are statistically distinguishable from zero.

2.2 Robustness Checks

We conducted a number of sensitivity checks to explore the robustness of our $s = 0$ results to lottery subsamples, household definition, and choice of estimator. The results from these analyses are summarized in Appendix Table C.1. The estimated effect of wealth on participation is similar across lotteries and results are robust to excluding spousal equity ownership from our participation definition. We also find that marginal effects from a probit estimator are substantively identical to the OLS estimates reported in the main text. Overall, our main findings are robust to these alternative specifications and analyses.
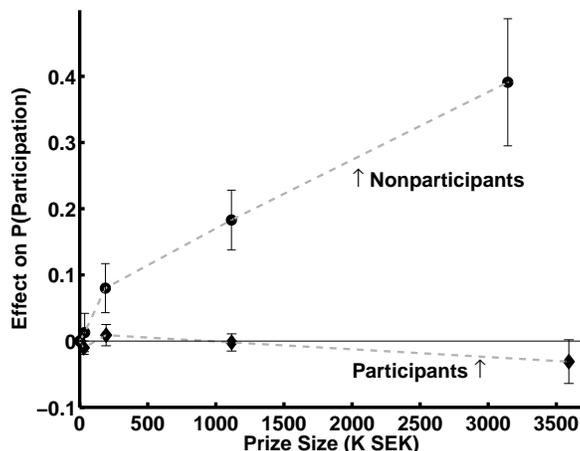
Figure 4: Effect of Wealth on Participation Probability by Prize Size. Coefficients are obtained by estimating Equation 1 in the post-1999 sample with the lottery wealth variable replaced by indicators for five mutually exclusive prize categories: 0 to 10K SEK (0 to 1.5K USD), 10K to 100K (1.5 to 15K), 100K to 1M (15 to 150K), 1M to 2M (150 to 300K), and 2M+ (300K+). Coefficient estimates and the 95% confidence bands are plotted at the mean prize in each category. See Appendix Table C.6 for the underlying estimates.

2.3   Calibrating Participation Costs to Match Causal Estimates

The causal effects we estimate can be used to test the quantitative predictions of models in which the source of non-participation is modest per-period financial costs. Vissing-Jørgensen (2003) noted modest per-period financial costs were not a plausible explanation for the wealthiest households but that they could explain the non-participation of low- and medium-wealth households. Indeed, we find modest wealth shocks induce some households to enter equity markets, suggesting modest per-period costs are a plausible explanation for some of the non-participation observed in low-wealth households. Quantitatively, however, our estimates suggest this deterrent needs to be supplemented with other disincentives for a substantial share of non-wealthy nonparticipants.

To get some intuition for this result, consider the finding, depicted in Figure 4, that a windfall gain greater than 2M SEK is estimated to increase participation probability by 39 percentage points in nonparticipants. Evaluating Equation 3 at 2M SEK implies that a cost of at least 47K SEK (7K USD) is needed to explain continued non-participation. This cost is likely an underestimate both because it does not take into account pre-lottery wealth and because all prizes in this category were in fact larger than 2M SEK (mean 3.1M SEK). Thus, the per-period costs necessary to explain non-participation are an order of magnitude larger than the median of the distribution of participation costs calculated in Section 1.4.

The estimated causal effects suggest certain structures of disincentives. First, the precisely estimated null-effects of wealth on continued participation suggest that disincentives to continued participation are modest. If not, increased participation benefits caused by the large wealth shock

16
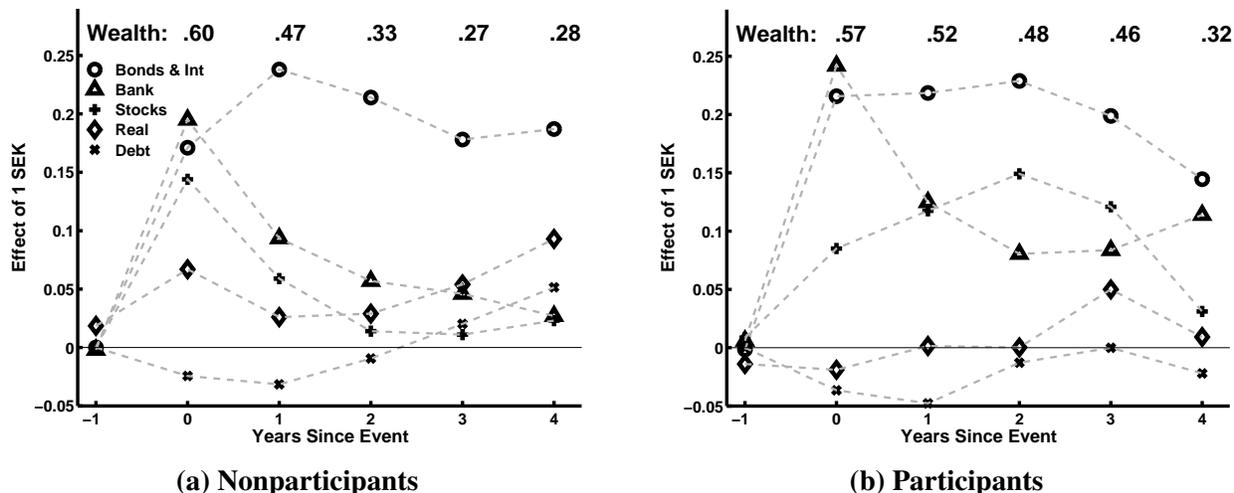
Figure 5: Effect of Wealth on Net Wealth, Bonds, Stocks, Bank Account Balances, Real Assets and Debt. Plotted are coefficients obtained from Equation 1 estimated in the post-1999 sample. For expositional clarity, the estimated effects on net wealth are displayed as numerical values. Bonds: interest rate funds, premium bonds and interest bonds. Stocks: directly and indirectly held stocks. Bank Acct: Bank Account Balances. Real Assets: property and land owned. Debt: gross debt. Net Wealth: financial and real assets minus debt. All outcome variables are valued at year-end market value. See Appendix Table C.7 for the underlying estimates at horizon $s = 0, 2, 4$.

would induce some participants who would otherwise exit to remain participants, thus generating a positive effect. Similarly, the finding that most pre-lottery nonparticipants do not enter following a large windfall gain suggests these households face large participation disincentives. Otherwise, a large wealth shock would induce more households to enter.

2.4 Where Does the Money Go?

Evidence on how participants and nonparticipants allocate the lottery wealth may provide important cues about how to interpret the costs calibrated in the previous section. Here, we use our main estimating equation to examine how lottery wealth affects five financial variables measured in the Swedish Wealth Register: bank account balances, stocks, bonds, debt and real assets. Figure 5 graphically depicts the main results from these analyses (see the accompanying caption for details on variable definitions). All variables are measured at year-end market value and all analyses are run in the post-1999 sample. The coefficients plotted for each variable are scaled so an estimate of 0.10 means that the total market value of the outcome (e.g., risky assets owned) increases by .1 for each 1 SEK won. Shown at the top of the figure are the numerical values of the coefficient estimates from an analogous specification in which the dependent variable is net wealth.

As expected, measured net wealth increases substantially following a windfall gain. In participants and nonparticipants alike, the register measure of net wealth increases in the year-of-win by roughly 60% of lottery prize won. This effect weakens over time, but remains at 30% at $s = 4$. For

17

a number of reasons, the $s = 0$ estimate of 60% does not imply that 40% of lottery wealth is consumed in the year-of-win. Statistics Sweden's net wealth measure will ordinarily not capture increases in the resale value of housing if some players use lottery wealth for home improvements, nor the resale value of cars and other durables purchased in the year of win. The net wealth variable also excludes monetary transfer to family members, unless the transfer is to the player's spouse. In exploratory analyses not shown here, we find that the effect of net wealth is approximately 9 percentage points higher if the wealth of unmarried, cohabiting partners are included.

The effects of lottery wealth on the net wealth trajectories of participants and nonparticipants are similar, and with the possible exception of stocks, as are the trajectories of the disaggregated variables. In both samples, players allocate 17-22% of wealth to bonds in the year-of-win, and this fraction is stable over the four-year period. Unsurprisingly, year-end bank account balances increase substantially in the year-of-win, but less than 12% of the lottery wealth remains in bank accounts at year-end in $s = 1$. Finally, Figure 5 shows that in both samples, the fraction of wealth allocated to purchase of new real estate or to pay off debts is modest.

Several of the results in Figure 5 are potentially relevant for interpreting the high costs reported in the calibration exercise of the previous section. First, the fact that less than 12% of the lottery wealth remains in bank accounts at $s = 1$, coupled with the observation that bond and equity investments take place shortly after the lottery, suggests a limited role for procrastination. Second, the small effects on debt and real assets suggest that few households forgo equity market participation to purchase new real estate or pay off high-interest debt. Third, nonparticipants and participants allocate more lottery wealth to bonds than equities. Fourth, among nonparticipants the effect of wealth on the market value of stocks owned is negligibly small from $s = 2$ onward. This supports the extensive margin finding that wealth shocks have small effect on nonparticipants' willingness to invest in equities. Overall, these findings suggest that many households evaluate the relative merits of bonds and equities very differently than standard theories predict.

In Figure 6, we show results from additional analyses of extensive-margin effects on bond market participation and real estate ownership. Panel (a) shows the estimated effect of wealth at $s = -1, ..., 4$ on binary indicators for three distinct states: (i) owns real estate but not equity, (ii) owns equity but not real estate, (iii) owns real estate and equity. The estimation sample is restricted to post-1999 households who owned neither real estate or stocks prior to the lottery. As noted above, the effect of lottery wealth on the value of real assets will be understated if some lottery wealth is invested in home improvements whose positive effect on the resale value of the property are understated in the register data. Since the sample does not include home owners, such concerns do not apply. Panel (a) shows that there is some evidence of delayed adjustment, but the estimated effect of 1M SEK on home ownership never exceeds 11 percentage points. Hence, though wealth causes some households to invest in housing, such behavior is not common. Figure

18
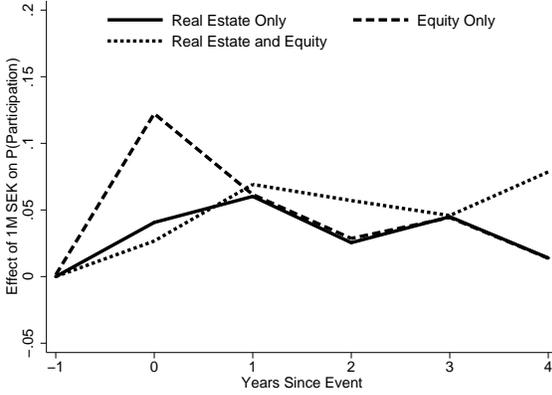
(a) Equity and Real Estate Market
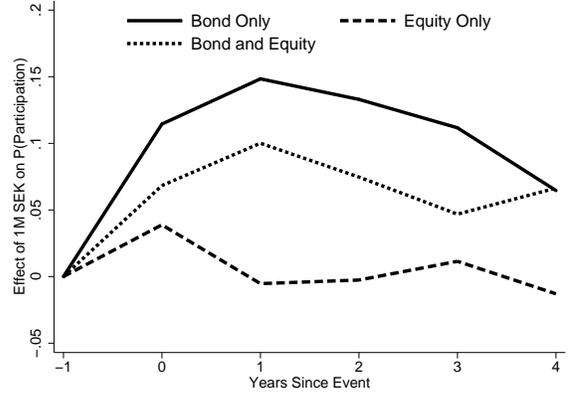
(b) Equity and Bond Market

Figure 6: Effect of Wealth (1M SEK) on Participation Probability in Equity, Real Estate, and Bond Markets. Panel (a) presents the effect of wealth on participation in real estate and equity markets for households in the post-1999 sample that participated in neither pre-lottery. Panel (b) presents the effect of wealth on participation in bond and equity markets for households in the post-1999 sample that participated in neither pre-lottery. See Appendix Tables C.8 and C.9 for the underlying estimates at horizons $s = 0, 2, 4$.

5 and Panel (a) of Figure 6 jointly suggest that real estate investments are unlikely to explain the relatively high rates of continued non-participation.

Panel (b) shows analogous estimates from an estimation sample restricted to households who owned neither bonds nor equities prior to the lottery. This sample overlaps greatly with our sample of pre-lottery nonparticipants because 90% of nonparticipants also did not own any bonds. The effect on bond market participation probability is more than twice as large as the effect on equity market participation probability at all horizons. Of the three possible states (bond only, equity only, bond and equity), we estimate the largest effects of wealth on the probability of participation in bond markets only. Figure 5 and Panel B Figure 6 thus suggest that among the households who did not participate in bond or stock markets pre-lottery, wealth is much more likely to trigger bond market entry than stock market entry.

3 A Structural Model

In this section we compare our causal estimates to the predictions of a plausibly calibrated life-cycle model of equity market participation. The model we study is standard, with a basic structure similar to leading studies such as Gomes and Michaelides (2005), Cocco (2005), Alan (2006), Khorunzhina (2013), and Fagereng et al. (2015). It is purposefully kept simple to highlight the link between our reduced-form findings and the model's predictions, and we regate detailed treatment of the household decision problem to Appendix A.

Briefly, agents in our model are finitely lived, facing an exogenous mortality probability and

19

have time separable, CRRA preferences with risk aversion of 4. Each period an agent optimally chooses how much to consume, save, and invest in equity markets given the agent's age, wealth, and participation status. Agents who choose to participate in equity markets face two separate types of costs. Participation costs, denoted κ, are paid in each period an agent allocates non-zero wealth to equity holdings. Entry costs, denoted χ, are paid whenever a previously non-participating agent decides to enter equity markets for the first time. Entry and participation costs are drawn independently from the two distributions:

$$(4) \qquad\qquad \kappa_i \sim F_{\theta_\kappa}(\kappa)$$
$$\chi_i \sim G_{\theta_\chi}(\chi),$$

where θ_κ and θ_χ are vectors that characterize the distribution of entry and participation costs. Equity provides a risky return r_s with $\mathbb{E}(r_s) > r_f$, where r_f is the risk-free rate. In addition, each period an agent is endowed with stochastic labor income y_t drawn from an age-specific distribution. Both income and equity returns are calibrated to match historical Swedish observations, yielding a 6.7% baseline equity premium and the income profile presented in Appendix Figure B.2.

3.1 Calibration

To compare the model's predictions to our causal estimates, we solve the baseline model, simulate a data set, and estimate the effect of lottery prizes on participation in the simulated data. Specifically, we draw lottery prizes identical to those in the data and assign these windfall gains to households with pre-lottery characteristics (including cells) identical to those in our post-1999 sample. We then solve for the optimal participation status of each household and estimate Equation 1. This permits a straightforward comparison of the model estimates to the causal estimates.

In our calibration exercises, we set entry costs χ_i equal to zero and assume κ_i is distributed according to the cumulative distribution function pictured in Figure 1. Table 6 compares the predicted effects of wealth (column 2) to the causal estimates (column 1) for the all-year sample (row 1) and the post-1999 sample stratified by participation status (rows 2 and 3). In the all-year sample, the model-generated coefficient is 0.140, roughly three times the magnitude of our causal estimate. For pre-lottery nonparticipants, the model-predicted effect is .224, again exceeding the causal estimate of .121. The number of nonparticipating households deterred from entering equity markets under our cost calibration is thus too small to match our causal estimates. For pre-lottery participants the model predicts, contrary to the causal estimates, a positive effect on continued participation. Intuitively, the calibrated costs cause some households with poor income or equity returns shocks to drop out of equity markets absent the windfall gain. In summary, the baseline calibration's predictions are generally qualitatively correct but overstate the effect of wealth on

Table 6: Comparison of Model-Predicted Effects of 1M SEK Windfall Gain to Lottery-Based Estimates. This table compares the model-implied coefficients for the effect of wealth and for the effect stratified by pre-lottery participation status under various calibration with the causal estimates (Column 1).

Effect	Causal Estimate (1)	Baseline (2)	High Per-Period Cost (3)	Correlated Income & Returns (4)	High Risk Aversion (5)	Lower Equity Premium (6)	High MPC (7)	All Disin-centives (8)
Baseline	.037	.140	.152	.129	.112	.093	.102	.071
Participants	-.001	.104	.117	.099	.093	.073	.075	.052
Nonparticipants	.121	.224	.232	.193	.169	.149	.163	.124

participation.

We also explored how well some standard extensions of the basic life-cycle framework improve the model's ability to replicate the causal estimates. The results of these analyses are summarized in Table 6. For point of reference, Column 1 shows our causal estimates, and Column 2 shows the effects predicted by the basic model. In Column 3 we consider the effects of multiplying each drawn per-period participation cost by two. Next, following Viceira (2001), we allow for a correlation of .25 between the random component of income and equity returns. Column 5 shows the results from a model calibrated with a risk-aversion parameter set to 10. Column 6 shows the predicted effects by a model which allows for the possibility that households may be unable to fully realize welfare gains of participation because of poor diversification and asset management fees (Calvet et al. (2007)). Following, Calvet et al. (2007), we impose a 1.5% management fee and scale the expected equity premium by 1/2, thus reducing the equity premium to 2.7%. Column 7 shows results obtained if the lottery windfall gain is rescaled by a factor of 0.6, an adjustment that may be appropriate if 40% of windfall gains are consumed immediately. Finally, Column 8 shows that when all extensions are included simultaneously, the model matches our causal estimate from the subsample of nonparticipants quite well, but continues to substantially overpredict the response in the full sample and in pre-lottery participants.

Although not exhaustive, these exercises suggest that standard models have difficulty matching the small effects of wealth on equity market participation under common calibrations. In interpreting these results, an important factor to consider is that we purposefully erred on the side of selecting parameter values that would increase the explanatory power of the extension. For example, a coefficient of risk aversion of 10 is the upper bound of the range considered plausible in Mehra and Prescott (1985) famous demonstration of the equity-premium puzzle, and for reasons explained in Section 2.4, 40% is a substantial overestimate of the actual propensity to consume out of the windfall gains.

(a) Entry	(b) Participation
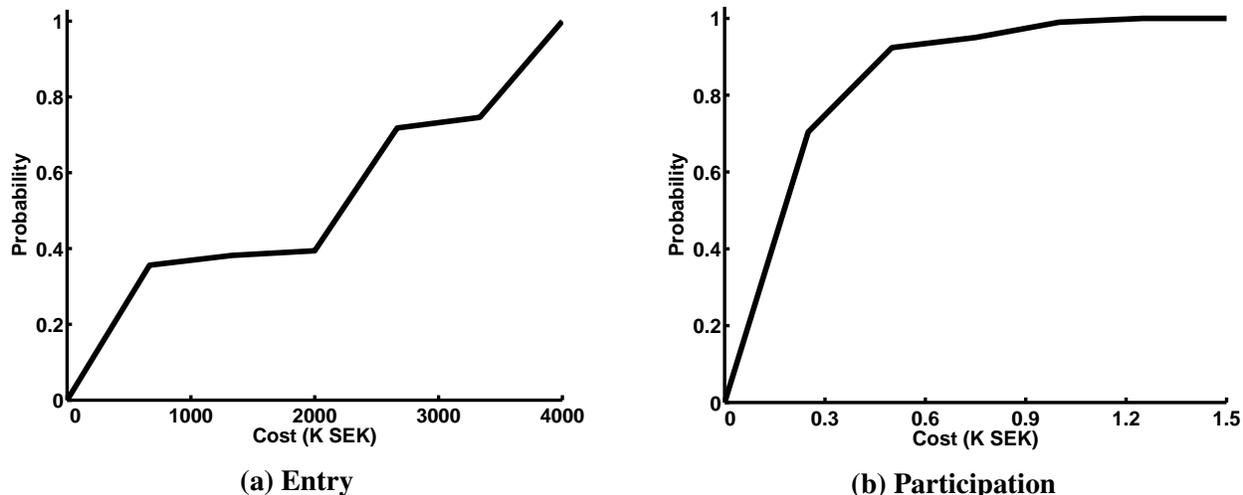
Figure 7: Structural Estimates of Fixed-Entry and Per-Period Participation Costs.

3.2 Estimation

To determine the size and structure of disincentives necessary for the model to match causal esti-mates, we use the method of indirect inference to estimate cost distributions given in Equation 4.[6] The resulting cost estimates can be interpreted as the financial equivalent of all disincentives nec-essary to generate the causal estimates. For additional details regarding the estimation procedure, including the resulting model fit, see Appendices A and B.

Figure 7 shows the estimated cost distributions. Fixed per-period costs of participation are estimated to be quite small, with a median estimated cost of 200 SEK (30 USD) and almost all households having estimated participation costs less than 1000 SEK (160 USD). Intuitively, small estimated per-period participation costs are required to match the reduced form finding that wealth has no discernible effect on continued participation. The estimated entry costs suggests that entry costs for nonparticipants are quite large, with 63% of entry costs estimated to be greater than 1M SEK (150K USD), 61% greater than 2M SEK (300K USD), and a 27% greater than 3M SEK (450K USD).[7] These large costs reflect the need for large disincentives to match the small effect of wealth on market entry, and thus confirm the qualitative patterns discussed in Section 2.3. The entry cost distribution is quite polarized, with only 2% of the entry costs estimated to be between 1M SEK (150K USD) and 2M SEK (300K USD). The 37% of households estimated to have entry costs below 1M SEK (150K USD) roughly correspond to those whose participation status responds to the range of wealth shocks we consider, while the remainder reflect households whose behavior

---

[6]Specifically, our estimation matches the causal coefficients from the all-year sample, the coefficients from the post-1999 sample stratified by pre-lottery participation status, and the coefficients from the non-linear specification considered in Figure 4. The exact coefficients matched are listed in Column 1 of Table A.1.

[7]The entry cost is modeled as a first-time cost of entry, and thus only affects decisions of pre-lottery nonparticipants.

is not well captured by the mechanisms included in our model.

Finally, we allow entry costs to vary by the pre-lottery characteristics considered in our heterogeneity analyses in Table 5. The estimated costs, shown in Appendix Table A.2, are intuitive given the reduced-form estimates, with costs estimated to be higher in subpopulations with smaller estimated causal effects. An important conclusion from this analysis is that the estimated entry costs exceed 1M SEK (150K USD) for the majority of households in all sub-populations considered, suggesting that the hypotheses advanced by Vissing-Jørgensen (2003) as possible explanations for the non-participation of rich households may prove relevant for understanding the behavior of nonparticipating households in a broader range of demographic groups.

4 Discussion

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)). In part the imperfect understanding reflects the reality that putting theoretical predictions to stringent tests, and quantifying the importance of the mechanisms inside the theories that generate nonparticipation, often requires difficult-to-attain datasets.

Our study contributes to the literature on participation by providing credible and statistically precise estimates of how large, positive, windfall gains impact equity market participation, and exploring quantitatively how our estimates compare to the predictions of a standard lifecycle model. We find that 1M SEK (150K USD) increases the probability of owning stocks by 12 percentage points among pre-lottery nonparticipants, but has no effect on pre-lottery participants. The effect is near-immediate, seemingly permanent, and heterogeneous in intuitive ways. For example, our finding of larger wealth effects in college-educated individuals is qualitatively consistent with explanations of nonparticipation that emphasize human capital, financial literacy and cognitive constraints (Calvet et al. (2007); Van Rooij, Lusardi and Alessie (2011); Grinblatt et al. (2011); Benjamin et al. (2013)).

A standard lifecycle model makes qualitative predictions that align well with the reduced-form evidence, but under a variety of calibrations and extensions predicts wealth effects larger than those we estimate. This overprediction is pervasive across demographic groups: in all subsamples considered in our heterogeneity analyses, the median entry cost needed to account for our causal estimates exceeds 800K SEK (120K USD). Housing, a general aversion to financial markets or procrastination are unlikely to play major roles in explaining why costs of this magnitude are needed for the model-generated predictions to match our experimental estimates. Lottery wealth only modestly impacts the value of real estate owned or the likelihood of owning a home. Nonparticipants who experience large windfall gains do enter financial markets, but more frequently bond than equity markets. Finally because this entry is near-immediate – little lottery wealth remains in

the winners' bank accounts a year after the lottery event – our results are not easily explained by procrastination either.

A parsimonious interpretation of our overall pattern of results is that a substantial fraction of nonparticipants are simply reluctant to engage with stocks as an asset class. Consistent with this interpretation is our finding that in nonparticipants lottery wealth has only a tiny effect on the total value of equities owned over medium-term horizons: we estimate that from $s = 2$ and onward, a 100-SEK increase in lottery wealth increases the value of the equity portfolio of a pre-lottery nonparticipant by a mere 2 SEK.

In summary, our study finds that exogenously "dumping" a large amount of wealth on households causes them to enter the stock market at rate that is puzzlingly low from the viewpoint of many standard economic frameworks. The mechanisms highlighted in these standard frameworks have considerable merit, but a complete theory of nonparticipation will likely require at least some households to have nonstandard preferences or beliefs that induces reluctance toward stock ownership. Such reluctance may generated by a multitude of nonstandard assumptions about beliefs or preferences, including loss aversion and narrow framing (Barberis et al. (2006)), disappointment aversion (Ang et al. (2005)), ambiguity aversion (Epstein and Schneider (2010); Campanale (2011); Dimmock, Kouwenberg, Mitchell and Peijnenburg (2015)) pessimistic beliefs about returns (Biais et al. (2010); Hurd, Van Rooij and Winter (2011)) fears that the "game is rigged" (Guiso et al. (2008)), or perceptions that equity markets are too complex (Vissing-Jørgensen (2003)). Our study's findings thus point to these areas as potentially productive areas of future study.

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.

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.

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.

Biais, Bruno, Peter Bossaerts, and Chester Spatt, "Equilibrium asset pricing and portfolio choice under asymmetric information," *Review of Financial Studies*, 2010, p. hhp113.

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 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, "Learning, ambiguity and life-cycle portfolio allocation," *Review of Economic Dynamics*, 2011, *14* (2), 339–367.

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

Carroll, Christopher D, "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.

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace, "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players," *Quarterly Journal of Economics*, 2015, *Forthcoming*.

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

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, Roy Kouwenberg, Olivia Mitchell, and Kim Peijnenburg, "Ambiguity Aversion and Household Portfolio Choice," *Journal of Financial Economics*, 2015, *Forthcoming*.

Epstein, Larry G and Martin Schneider, "Ambiguity and asset markets," *NBER Working Paper No. 16181*, 2010.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso, "Asset market participation and portfolio choice over the life-cycle," *Journal of Finance*, 2015, *Forthcoming*.

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

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

Graham, John R and Campbell R Harvey, "Expectations of equity risk premia, volatility and asymmetry from a corporate finance perspective," *NBER Working Paper No. 8678*, 2001.

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**, "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, Maarten Van Rooij, and Joachim Winter, "Stock market expectations of Dutch households," *Journal of Applied Econometrics*, 2011, *26* (3), 416–436.

Imbens, GW, D Rubin, and B 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.

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

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

Mehra, Rajnish and Edward C Prescott, "The equity premium: A puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

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.

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.

Rooij, Maarten Van, Annamaria Lusardi, and Rob Alessie, "Financial literacy and stock market participation," *Journal of Financial Economics*, 2011, *101* (2), 449–472.

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

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, "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 2002, *110* (4), 825–853.

_ , "Perspectives on Behavioral Finance: Does" Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," in "NBER Macroeconomics Annual 2003, Volume 18," The MIT Press, 2003, pp. 139–208.

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

A Model

Each period, an age t agent chooses how much to consume, save, and invest in equity markets. An agent has a maximum lifespan of $T = 113$, but prior to reaching age T, faces mortality risk with an exogenous probability of surviving from period t to $t + 1$ denoted by s_t. Upon death, an agent receives a terminal payout of zero. Agents who invest in equity face two separate types of costs. Participation costs, denoted κ, are paid in each period an agent allocates non-zero wealth to equity holdings. Entry costs, denoted χ, are paid whenever a previously non-participating agent decides to enter equity markets for the first time. Entry and participation costs are drawn independently from the two distributions:

$$\kappa_i \sim F_{\theta_\kappa}(\kappa)$$

(5)

$$\chi_i \sim G_{\theta_\chi}(\chi),$$

where θ_κ and θ_χ are vectors that characterize the distribution of entry and participation costs. Equity provides a risky return r_s with $\mathbb{E}(r_s) > r_f$, where r_f is the risk-free rate. In addition, each period an agent is endowed with stochastic labor income y_t drawn from an age-specific distribution.

In each period, an age t agent chooses how much to consume, save, and invest in equity markets. For an agent who decides not to participate in equity markets in a given period, the decision problem reduces to deciding how much to consume and how much to save in the risk-free asset:

$$
\begin{aligned}
(6) \quad V_t^{NP}(W_t, \kappa, \chi) = \max_{c_t, W_{t+1}} & \ u(c_t) + \beta s_t \mathbb{E}_{y_{t+1}} \left[V_{t+1}(W_{t+1}, I_t, \kappa, \chi) \right] \\
& W_{t+1} = r_f \left(W_t - c_t \right) + y_{t+1} \\
& I_t = 1,
\end{aligned}
$$

where I_t is an indicator equal to 0 in year t if the agent has never participated in equity markets, and 1 otherwise. An agent who participates in the equity market, decides how much to consume, and what fraction of wealth to allocate to stocks. This decision problem can similarly be expressed as:

$$
\begin{aligned}
V_t^{P}(W_t, \kappa, \chi) = \max_{c_t, W_{t+1}, \alpha_t} & \ u(c_t) + \beta s_t \mathbb{E}_{y_{t+1}, r_{s,t+1}} \left[V_{t+1}(W_{t+1}, I_t, \kappa, \chi) \right] \\
(7) \quad & W_{t+1} = r_f \left(W_t - c_t - \kappa \right) + \alpha_t (r_{s,t+1} - r_f) \left(W_t - c_t - \kappa \right) + y_{t+1} \\
& I_t = 0.
\end{aligned}
$$

Subtracting the entry costs if the agent has not previously participated in equity markets yields

the following decision problem:

(8) $$V_t(W_t, I_{t-1}, \kappa, \chi) = \max\{V_t^{NP}(W_t, \kappa, \chi) \,, \; V_t^P(W_t - \chi I_{t-1}, \kappa, \chi)\}.$$

Each period, comparing the value function of non-participation against the value function of participation net any entry costs determines the agent's participation decision. We assume s_t, r_s and y_t follow known stochastic processes. Although our data are collected at the household level, we make the simplifying assumption s_t as the survival probability as a function of an individual at age t. We assume the income process is age-dependent, with log income of an age-t individual evolving according to the following process:

$$\ln y_t = f(t) + \sigma_{y,t}\eta_t,$$

where $f(t)$ is a quadratic in age and $\sigma_{y,t}$ is the age-specific standard deviation in earnings. Both the earnings and mortality processes are estimated using a cross section of the Swedish 1999 population. To match historical equity returns in Sweden (see Waldenström (2014)), we assume that equity returns are lognormally distributed, with location parameter set to 0.065 and scale parameter set to 0.21. We calibrate $r_f = 2\%$ to match historical yield on Swedish government bonds.

Structural Estimation Results – Table A.1 presents the model fit for the baseline structural estimation (see Figure 7 for the costs).

Estimation with Heterogeneous Costs – To understand how participation disincentives vary among subpopulations, we repeat our structural estimation, now allowing cost distributions to vary with characteristics included in Table 5. To identify these differences, we augment our baseline coefficient target $\hat{\beta}$ to include coefficients presented in Table 5, and re-estimate Equation 9, thus matching the heterogeneity in the effect of wealth on entry.

In implementing this estimation, we make two additional assumptions. First, because continued participation effects are negligible and no evidence of heterogeneity in the effect on continued participation exists, we do not allow for heterogeneity in per period costs of participation. Instead, for all households, we calibrate κ to 200 SEK (30 USD), roughly the median per-period participation cost in our baseline structural estimation. Second, to maintain computational feasibility, we assume a parametric cost distribution. Specifically, we assume costs of entry χ_i are distributed normally with truncation at zero to ensure non-negativity. We assume mean $\mu_\chi x_i$, where x_i is a vector of indicator variables corresponding to the characteristics in Table 5 (plus a constant) and μ_χ is a coefficient vector to be estimated.

Table A.1: Structural Estimation Results, Model Fit. Column 1 presents causal estimates while Column 2 presents the model-implied coefficients when using the estimated cost distributions.

Effect	Causal Estimates (1)	Model-Implied Estimates (2)
Baseline	.037	.040
Participants (NP)	-.001	.010
Nonparticipants (P)	.121	.129
15K-100K (P)	-.010	.009
100K-1M (P)	-.009	.022
1M-2M (P)	-.002	.008
2M+ (P)	-.031	.019
15K-100K (NP)	.013	.015
100K-1M (NP)	.080	.069
1M-2M (NP)	.183	.175
2M+ (NP)	.391	.401

Table A.2 presents resulting model fit and estimates, while the estimated CDFs of entry costs are presented in Figure A.1. These CDFs again highlight the extremely high costs for all subsamples, because no group of nonparticipants is estimated to have a median entry cost of less than 1M SEK (150K USD). Although cost patterns vary intuitively along various characteristics, in all groups we find that large disincentives to entry are necessary to replicate estimated effects. These patterns directly track the reduced-form estimates.
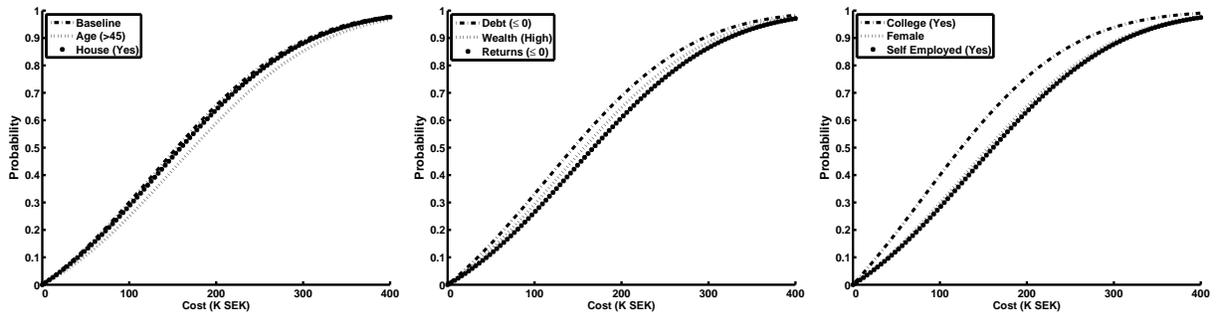
Figure A.1: Heterogeneous Structural Estimates of Entry Cost Distributions. For each distribution the mean is shifted from the baseline by the specifice characteristic estimate.

Table A.2: Heterogeneous Structural Estimation Results. Panel A presents the model estimated parameters. Specifically, the first row presents the estimated constant mean and standard deviation. The rows beneath present the estimated shift in the mean for an individual with the specified characteristic. Panel B presents the resulting model fit, with Column 1 presenting the causal estimates and Column 2 presents the model-implied coefficients when using the estimated cost distributions.

A. Parameter Estimates

μ_χ		σ_χ	
1,283,679		1,295,843	
Debt (≤ 0)	**Home Owner (Yes)**	**Wealth (High)**	**Equity Returns (≤ 0)**
-147,498	66,566	36,073	183,931
Sex (Female)	**Age (> 45)**	**College (Yes)**	**Self-Employed (Yes)**
15,434	262,141	-479,789	93,758

B. Model Fit

Characteristic		Causal Estimates (1)	Model-Implied Estimates (2)
Baseline		.037	.044
Participants (P)		-.001	.011
Nonparticipants (NP)		.121	.121
Prize Size	**15K-100K (P)**	-.010	.002
	100K-1M (P)	-.009	.028
	1M-2M (P)	-.002	.015
	2M+ (P)	-.031	.011
	15K-100K (NP)	.013	-.001
	100K-1M (NP)	.080	.060
	1M-2M (NP)	.183	.153
	2M+ (NP)	.391	.387
Debt	≤ 0	.212	.183
	> 0	.094	.074
Home Owner	**Yes**	.105	.086
	No	.146	.159
Wealth	**High**	.062	.072
	Low	.138	.130
Prior Equity Returns	> 0	.139	.124
	≤ 0	.056	.071
Sex	**Female**	.117	.123
	Male	.149	.131
Age	> 45	.113	.096
	≤ 45	.141	.127
College Educated	**Yes**	.232	.227
	No	.097	.097
Self-Employed	**Yes**	.023	.048
	No	.134	.122

B   Model Solution and Estimation

Model Solution and Details

The model solution algorithm utilizes a modified endogenous grid method algorithm. The endogenous grid is defined over the savings problem, while a grid search (100 grid points) is used to identify optimal portfolio choice (see Carroll (2006) and Barillas and Fernández-Villaverde (2007)) for more details. Model solution algorithm summary and model code are both available on request.

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. Note that there is no attrition or selection concerns in this sample as it is drawn randomly from the entire population. Although simple, we view it as unlikely that allowing for richer survival probabilities would substantially affect our qualitative conclusions. The resulting estimates are presented in Figure B.1.
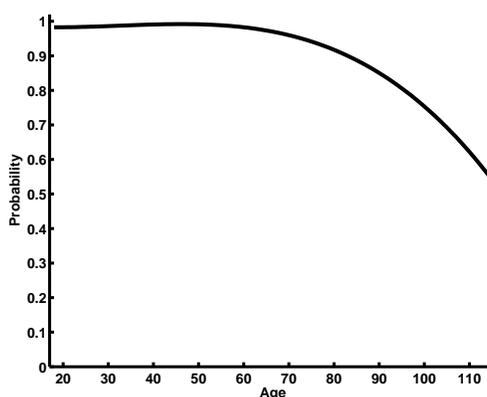
Figure B.1: Survival Probabilities. This figure presents the estimated one year survival probability for each age.

Income is calculated using the observed incomes in 1999. We select 100,000 individuals in year 1999 from the Swedish population, and combine with partners to get a measure of after tax household income. We then regress a quartic in age on this variable. We do not permit time, cohort, or sex effects in our estimation. Using our estimated income process, we then calculate the residual for each household. We then regress a quartic in age on the squared residual to obtain an age specific standard deviation of income realizations. In column 4 of Table 6 we consider the effect of allowing for correlation between income risk and equity returns. The resulting income estimates are presented in Figure B.2.
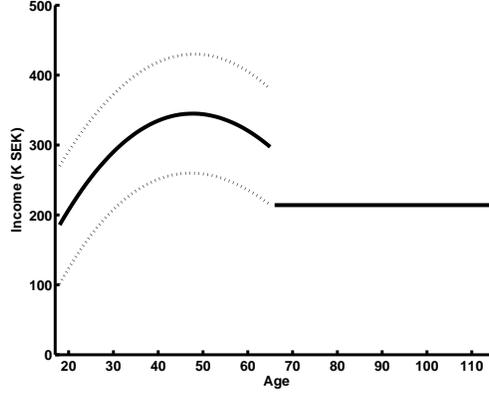
Figure B.2: Average Income Profiles. For ages below retirement, we also present the high and low income states.

Estimation Technicals

We estimate the distribution of individual costs of cost of entry and participation that most closely replicate the empirical estimates presented in column 1 of Table A.1. As shown in equation 4, these cost distributions are parametrized by vectors θ_κ, and θ_χ. Defining $\Theta = (\theta_\kappa, \theta_\chi)$, we estimate

$$(9) \qquad \hat{\Theta} = \arg\min_{\Theta}(\hat{\beta} - \tilde{\beta}(\Theta))'W(\hat{\beta} - \tilde{\beta}(\Theta))$$

where $\hat{\beta}$ represents the the vector of empirical estimates and $\tilde{\beta}(\Theta)$ represents the vector of coefficients implied by the model. W could be any positive semi-definite matrix, but in all estimations we use the identity matrix.

The cost distributions are estimated non-parametrically by defining θ_κ, and θ_χ as points on the CDF. We first bound the set of feasible costs by assuming $F_{\theta_\kappa}(\overline{\kappa}) = 1$ and $G_{\theta_\chi}(\overline{\chi}) = 1$. Assuming $F_{\theta_\kappa}(0) = 0$ and $G_{\theta_\chi}(0) = 0$, we construct a linearly spaced grid consisting of 7 points denoted as x_κ^n and x_χ^n, respectively. We assume the CDFs are piecewise linear between these points and define

$$\theta_\kappa = \{\theta_\kappa^n | F_{\theta_\kappa}(x_\kappa^n) = \theta_\kappa^n\}$$
$$\theta_\chi = \{\theta_\chi^n | G_{\theta_\chi}(x_\chi^n) = \theta_\chi^n\}.$$

The cost distributions are thus characterized by the estimates θ^n corresponding to the evaluation of the CDF at each point x^n. $\overline{\kappa}$ and $\overline{\chi}$ are not known prior to estimation, but are chosen such that the next-to-last point on each respective grid is reasonably close but not equal to 1.

To simulate the model and generate coefficients $\tilde{\beta}(\theta)$, we first solve the model over a grid of participation and entry costs and then proceed in several steps:

1. Take parameter set Θ as given.

2. From the time $s = -1$ data set sample for all lottery players all relevant variables, including observed wealth, participation decisions, and time invariant covariates included in Table 4. From the time $s = 0$ data set, sample the amount won.

3. For each sampled observation, draw two uniform random variables. Using CDFs parametrized by Θ, map these random variable to a realized (κ_i, χ_i).

4. According to (κ_i, χ_i), interpolate over grid of participation and entry costs to calculate for optimal decision rules.

5. Taking the time $s = -1$ data set as given, assign lottery wealth and calculate the optimal $s = 0$ saving and participation decisions.

6. Repeat reduced form estimation on the model simulated data set to obtain coefficients $\tilde{\beta}(\theta)$.

Following each of these calculations the objective function presented in Equation 9 is calculated. A numerical optimization algorithm is utilized to find the global minimum of this objective function.

C Supplemental Tables

Table C.1: Effect of Wealth (1M SEK) on Participation Probability, Robustness Checks. This table presents the estimated effects of wealth at time $s = 0$ for various lottery subsamples, when the household definition is restricted to individual winners, and for a probit estimator. 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 Subsamples | | | | | |
| | Kombi | | | Triss | | |
	Pooled	P	NP	Pooled	P	NP
Effect	.035	.002	.156	.036	-.004	.104
SE	(.018)	(.008)	(.037)	(.012)	(.006)	(.030)
N	31,903	21,306	7,972	1,969	1,070	710

| | PLS | | |
	Pooled	P	NP
Effect	.143	-.005	1.99
SE	(.100)	(.013)	(2.90)
N	45,905	30,994	11,606

| | Other Robustness | | | | | |
| | Probit | | | Individual Analysis | | |
	Pooled	P	NP	Pooled	P	NP
Effect	.114	-.014	.500	.042	.004	.100
SE	(.036)	(.066)	(.091)	(.011)	(.004)	(.021)
Marg. Effect	.037	-.001	.104	—	—	—
N	79,777	51,323	19,403	79,777	49,012	24,646

Table C.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 near-identical to that used by by Calvet et al. (2007). Marginal effects are calculcated 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**	**Change**	**Estimate**	**SE**	**Change**
	(1)	(2)	(3)	(4)	(5)	(6)
Disposable Income (100K SEK)	.363	.049	3.45%	.450	.096	3.31%
Financial Assets (log)	.206	.004	39.59%	.180	.004	22.12 %
Total Real Estate (log)	.025	.006	11.45%	.028	.006	9.38%
Total Liabilities (log)	-.009	.006	-.471%	-.014	.006	-5.99 %
Retired	.074	.029	1.43%	.060	.026	.89%
Self Employed	.011	.041	.07%	.021	.032	.13%
Unemployed	.084	.027	.94%	.033	.027	.26%
Student	.099	.043	.72%	.164	.047	.83%
Age	-.012	.001	-8.04%	-.010	.001	-4.69%
Household Size	.002	.010	.07%	.014	.009	.40%
High School	.171	.021	1.77%	.124	.018	1.73%
Higher Degree	.267	.022	2.73%	.251	.020	3.31%
Missing Education	.189	.083	.33%	.550	.101	1.42%
Immigrant	-.096	.044	-.62%	-.281	.031	-2.22%
Constant	-4.361	.116		-3.69	.096	
N	73,644			49,959		

Table C.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 ($\hat{y}|L_i = 0$) 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	.003	.012	73,658	.725	.0002	.012	73,658	.423
0	.037	.010	79,777	.725	.020	.011	79,777	.430
1	.046	.010	97,782	.722	.037	.011	97,782	.429
2	.047	.010	121,134	.738	.044	.012	121,134	.454
3	.047	.009	150,808	.750	.045	.011	150,808	.470
4	.045	.010	159,974	.758	.045	.012	159,974	.484
5	.045	.010	165,254	.759	.033	.011	165,254	.490
6	.043	.010	181,785	.764	.045	.012	181,785	.499
7	.036	.010	198,122	.773	.047	.012	198,122	.510
8	.052	.010	214,931	.779	.051	.012	214,931	.520
9	.052	.010	205,641	.782	.049	.012	205,641	.528
10	.043	.011	182,411	.778	.028	.013	182,411	.530

Table C.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 ($\hat{y}|L_i = 0$) from estimating Equation 1 in the post-1999 sample. Columns 9 through 16 show analogous estimates with participation defined more narrowly to only include directly owned stocks.

	A. Stock or Mutual Fund								**B. Stock Only**							
Horizon	**P**				**NP**				**P**				**NP**			
(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)**
-1			55,897				21,812				32,623				45,086	
0	-.001	.005	53,370	.978	.121	.025	20,288	.069	.009	.003	31,181	.973	.018	.010	42,477	.040
1	.002	.010	50,354	.955	.114	.028	18,664	.092	-.001	.014	29,424	.950	.021	.011	39,594	.050
2	.013	.009	46,455	.938	.089	.031	16,822	.114	.012	.012	27,199	.927	.024	.013	36,078	.060
3	.021	.008	43,069	.922	.071	.032	15,101	.129	.014	.018	25,304	.905	.025	.016	32,866	.066
4	.019	.010	39,822	.906	.078	.043	13,565	.140	-.006	.029	23,474	.885	.032	.021	29,913	.072

Table C.5: Heterogeneous Effect of Wealth (1M SEK) on Participation Probability. Coefficients are obtained by estimating Equation 1 in the post-1999 sample of nonparticipants at time $s = 3$. 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.

	Financial Characteristics							
	Debt		Home Owner		Net Wealth		Equity Returns	
	≤ 0	> 0	Yes	No	High	Low	≤ 0	> 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Effect	.185	.055	.064	.072	.059	.125	.052	.095
SE	(.107)	(.032)	(.045)	(.046)	(.033)	(.076)	(.040)	(.053)
p	.084	.080	.159	.116	.076	.105	.193	.072
Hetero p	.245		.899		.428		.521	
N	7,926	7,175	7,872	7,229	3,666	11,435	9,869	5,232

	Demographic Characteristics							
	Sex		Age		College		Self-Employed	
	Male	Female	≤ 45	> 45	Yes	No	Yes	No
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Effect	.163	.038	.060	.070	.098	.059	.072	.124
SE	(.053)	(.036)	(.035)	(.056)	(.063)	(.034)	(.032)	(.087)
p	.002	.284	.085	.211	.122	.085	.028	.155
Hetero p	.052		.886		.590		.575	
N	6,626	8,475	2,289	12,812	3,856	11,245	455	14,646

Table C.6: 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 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. In Triss, all prizes exceed 1.5K, so the omitted category is prizes below 15K USD.

A. Pooled Sample

	P				NP			
	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate	-.010	-.009	-.002	-.031	.013	.080	.183	.392
SE	(.009)	(.016)	(.013)	(.033)	(.029)	(.037)	(.045)	(.096)
ME	-.342	-.045	-.002	-.009	.357	.423	.164	.125
N	485	804	204	50	257	525	94	28

B.Kombi

	P				NP			
	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate			-.003	.042			.187	.271
SE			(.013)	(.009)			(.047)	(.188)
ME			-.003	.014			.168	.110
N			190	10			84	6

C. Triss

	P				NP			
		$15K < L_i \leq 150K$	$1M < L_i \leq 2M$	$2M < L_i$		$15K < L_i \leq 150K$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate		.001	.032	-.043		.079	.155	.417
SE		(.015)	(.016)	(.042)		(.021)	(.136)	(.103)
ME		.006	.029	-.011		.417	.141	.125
N		804	12	40		525	10	22

D. PLS

	P				NP			
	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$	$10K < L_i \leq 100K$	$100K < L_i \leq 1M$	$1M < L_i \leq 2M$	$2M < L_i$
Estimate	-.012		.007		.019			
SE	(.010)		(.002)		(.029)			
ME	-1.05		.006		1.68			
N	271		2		104			

Table C.7: Effect of Wealth on Net Wealth, Bonds, Stocks, Bank Account Balances, Real Assets and Debt. This table presents results from estimating Equation 1 in the all-year (pooled) and post-1999 (P and NP) samples. We present results for the full sample and subsamples of participants and nonparticipants at horizons $s = 0, 2, 4$. Lottery wealth and financial variables are measured in SEK.

		Total Wealth	Real Assets	Financial Assets			Debt
				Risky	Safe	Bank	
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Pooled	$s = 0$	.611	.029	.139	.188	.218	-.023
	N=79,777	(.044)	(.031)	(.033)	(.027)	(.027)	(.017)
	$s = 2$	.486	.007	.161	.197	.090	-.009
	N=165,254	(.045)	(.022)	(.035)	(.023)	(.023)	(.017)
	$s = 4$	.393	.031	.117	.153	.071	.002
	N=182,411	(.046)	(.021)	(.029)	(.016)	(.017)	(.017)
P	$s = 0$	.572	-.020	.085	.215	.241	-.037
	N=53,370	(.035)	(.011)	(.022)	(.021)	(.019)	(.011)
	$s = 2$	.483	.001	.149	.228	.080	-.013
	N=43,069	(.064)	(.058)	(.057)	(.044)	(.025)	(.030)
	$s = 4$	.319	.009	.031	.144	.113	-.023
	N=20,713	(.069)	(.054)	(.030)	(.027)	(.046)	(.017)
NP	$s = 0$	.602	.067	.144	.170	.195	-.025
	N=20,288	(.061)	(.034)	(.060)	(.047)	(.033)	(.019)
	$s = 2$	.326	.029	.014	.214	.056	-.010
	N=15,101	(.053)	(.022)	(.008)	(.048)	(.016)	(.019)
	$s = 4$	.277	.093	.023	.186	.026	.051
	N=7,002	(.089)	(.045)	(.011)	(.059)	(.012)	(.027)

Table C.8: Effect of Wealth (1M SEK) on Real Estate and Equity Market Participation Probability.
This table presents the effects of wealth on real estate and equity market participation probability. Each number in a given row represents the effect of winning 1M SEK on the probability that a household enters the state specified in a column. This is repeated at horizon $s = 0, \ 2, \ 4$. In all estimations controls in Table 4 are included.

Time s=-1 (Real Estate/Equity)	**A. At s=0 (Real Estate/Equity)**			
	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
(N/N) ($N = 8,469$)	-.190 (.038)	.041 (.030)	.122 (.044)	.027 (.019)
(Y/N) ($N = 11,819$)	-.009 (.005)	-.097 (.027)	.002 (.004)	.103 (.027)
(N/Y) ($N = 10,402$)	-.013 (.013)	.014 (.010)	-.056 (.039)	.054 (.035)
(Y/Y) ($N = 42,968$)	.004 (.002)	-.001 (.005)	.009 (.007)	-.011 (.009)

s=-1 Own (Real Estate/Equity)	**B. At s=2 Own (Real Estate/Equity)**			
	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
(N/N) ($N = 7,027$)	-.111 (.058)	.025 (.038)	.029 (.025)	.057 (.034)
(Y/N) ($N = 9,795$)	-.008 (.011)	-.081 (.041)	-.003 (.006)	.091 (.043)
(N/Y) ($N = 9,186$)	.030 (.023)	-.021 (.013)	-.101 (.049)	.090 (.045)
(Y/Y) ($N = 37,269$)	-.001 (.002)	-.013 (.010)	.015 (.012)	-.001 (.015)

s=-1 Own (Real Estate/Equity)	**C. At s=4 Own (Real Estate/Equity)**			
	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
(N/N) ($N = 5,528$)	-.106 (.078)	.014 (.040)	.014 (.026)	.070 (.059)
(Y/N) ($N = 8,037$)	-.023 (.015)	-.042 (.059)	-.006 (.003)	.072 (.064)
(N/Y) ($N = 7,817$)	.026 (.048)	-.0003 (.011)	-.093 (.077)	.067 (.061)
(Y/Y) ($N = 32,005$)	-.007 (.002)	-.020 (.009)	-.001 (.007)	.028 (.012)

Table C.9: Effect of Wealth (1M SEK) on Bond and Equity Market Participation Probability. This table presents the effects of wealth on bond and equity market participation probability. Each number in a given row represents the effect of winning 1M SEK on the probability that a household enters the state specified in a column. This is repeated at horizon $s = 0,\ 2,\ 4$. In all estimations controls in Table 4 are included.

Time s=-1 (Bonds/Equity)	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
	A. At s=0 **(Bonds/Equity)**			
(N/N) ($N = 17,510$)	-.222 (.035)	.114 (.026)	.039 (.021)	.068 (.020)
(Y/N) ($N = 2,718$)	-.005 (.064)	-.175 (.069)	.064 (.046)	.117 (.073)
(N/Y) ($N = 33,677$)	.001 (.007)	.006 (.004)	-.185 (.029)	.178 (.027)
(Y/Y) ($N = 19,693$)	-.002 (.001)	-.006 (.002)	-.023 (.008)	.031 (.008)

s=-1 Own (Bonds/Equity)	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
	B. At s=2 Own **(Bonds/Equity)**			
(N/N) ($N = 14,514$)	-.205 (.044)	.133 (.033)	-.003 (.018)	.075 (.024)
(Y/N) ($N = 2,308$)	-.128 (.056)	-.040 (.092)	.020 (.044)	.148 (.081)
(N/Y) ($N = 29,223$)	-.020 (.013)	.015 (.008)	-.208 (.033)	.213 (.035)
(Y/Y) ($N = 17,232$)	-.008 (.002)	-.008 (.006)	-.012 (.027)	.028 (.028)

s=-1 Own (Bonds/Equity)	(N/N) (1)	(Y/N) (2)	(N/Y) (3)	(Y/Y) (4)
	C. At s=4 Own **(Bonds/Equity)**			
(N/N) ($N = 11,633$)	-.118 (.057)	.065 (.032)	-.013 (.025)	.066 (.029)
(Y/N) ($N = 1,932$)	-.193 (.074)	-.220 (.153)	.004 (.150)	.409 (.208)
(N/Y) ($N = 24,847$)	-.028 (.014)	.016 (.012)	-.147 (.047)	.158 (.045)
(Y/Y) ($N = 14,975$)	-.012 (.004)	-.005 (.009)	-.043 (.037)	.059 (.039)