Wealth and Stock Market Participation: Estimating the Causal Effect From Swedish Lotteries *

Joseph Briggs† David Cesarini‡ Erik Lindqvist § Robert Östling¶

Preliminary
May 3, 2015

Abstract

The positive cross-sectional relationship between wealth and equity market participation suggests that moderate participation costs can rationalize most participation decisions. This study uses random assignment of over 500 million USD to Swedish households to precisely identify the causal effect of wealth on participation. The effect is estimated to be positive, but much smaller than that implied by the cross-section. Structural estimates of the costs necessary to explain the estimated effects of wealth on participation are too high to be credible, even after permitting cost heterogeneity conditional on individual characteristics. It is thus unlikely that fixed financial costs are the cause of equity market non-participation. Reinterpreting the disincentive to participate as pessimistic subjective beliefs of equity returns results in a belief distribution that is both credible and consistent with prior measurement of beliefs.

1 Introduction

Canonical life-cycle models of consumption and savings predict that all individuals, irrespective of their degree of risk aversion, should invest some non-zero fraction of their wealth in stocks (Samuelson (1969); Merton (1971)). Because this prediction is not borne out empirically – a substantial fraction of household do not own stock directly or through mutual funds (Friend and Blume (1975); King and Leape (1984); Mankiw and Zeldes (1991)) – a large literature in household finance formulates and tests hypotheses about the causes of the "non-participation puzzle" (Haliassos and Bertaut (1995)). Insights into the causes of the non-participation puzzle may lead to the formulation of better models of household finance and may also help guide efforts to promote more effective financial decision making by households (Campbell (2006)).

To account for the high rates of non-participation, many models introduce a fixed cost of stock ownership, either a one-time entry cost or an ongoing, per-period, participation cost. Because the potential gains increase with greater wealth, whereas the costs do not, these models predict that a

*This paper is part of a project hosted by the Research Institute of Industrial Economics (IFN). We are grateful to IFN Director Magnus Henrekson for his strong commitment to the project and to Marta Benkestock for superb administrative assistance. We also thank Claudio Campanale for very helpful comments. The project is financially supported by two large grants from the Swedish Research Council (VR) and Handelsbanken's Research Foundations. We also gratefully acknowledge financial support from the NBER Household Finance working group, US National Science Foundation and the Swedish Council for Working Life, and Social Research (FAS).

†Corresponding author. Joseph Briggs, Department of Economics, New York University, Research Institute for Industrial Economics. Email: jsb493@nyu.edu.

‡David Cesarini, Department of Economics, New York University, Research Institute for Industrial Economics. Email: david.cesarini@nyu.edu

§Erik Lindqvist, Stockholm School of Economics, Research Institute for Industrial Economics. Email: erik.lindqvist@hhs.se

¶Robert Östling, Institute for International Economic Studies. Email: robert.ostling@iies.su.se

household will enter the stock market if their wealth exceeds an endogenously determined threshold. The models therefore predict that a wealth shock should increase the probability of participation, and thus provide an intuitive explaination for the robustly established positive cross-sectional correlation between wealth and participation (Mankiw and Zeldes (1991); Poterba and Samwick (2003); Calvet, Campbell, and Sodini (2007)). To explain this relationship, Vissing-Jorgensen (2002) showed that a very modest per-period cost can account for observed non-participation, a finding that has been confirmed in subsequent studies (see Section XX).

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. Theories of household finance make predictions about the impact of a windfall gain on subsequent stock market participation and estimates of the causal impact of wealth on participation are therefore useful for testing and refining theories of household finance. A fundamental challenge when estimating the effect of wealth is that it is not feasible to randomly assign substantial amounts of wealth to individuals. As a result, researchers interested in the relationship between wealth and participation are usually confined to studying observational data (e.g., Calvet and Sodini (2014); Brunnermeier and Nagel (2008); Calvet, Campbell, and Sodini (2009)), where the possibility of omitted variable bias and reverse causation looms large (though see Andersen and Nielsen (2011) for an example of a quasi-experimental study exploiting unexpected parental deaths as a plausibly exogenous source of wealth variation). The overall conclusion from this literature is that changes in wealth are associated with a greater likelihood of participation, but the magnitudes of the estimated effects vary. A second branch of the literature (e.g., Gomes and Michaelides (2005); Alan (2006); Khorunzhina (2013); Fagereng, Gottlieb, and Guiso (2013)) uses structural models to identify the magnitudes and patterns of such costs, with most finding that moderate financial costs are capable of accounting for the majority of non-participation.

Our paper contributes to both of these literatures. First, we conduct reduced form analyses and report a comprehensive set of analyses examining how wealth impacts participation. Our sample satisfies a number of methodological desiderata that allow us to make stronger inferences about the effect of wealth on participation than in previous work. First, we observe the factors (such as number of tickets owned) conditional on which the lottery wealth is randomly assigned. We show that randomozation checks are passed and we can be uniquely confident that our estimates have a causal interpretation. 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, and any sample selection biases should be negligibly small. Our reduced form estimates suggest that wealth has a small, significant effect on participation. Furthermore, we estimate significant heterogeneity in the effect amongst individuals with various characeteristics, although generally the estimated effects remain small.

In our structural analyses, we use the exogenous wealth variation to estimate a structural model of portfolio choice over the lifecycle. To convey the intuition behind the principal result in our structural analyses, consider again the result that a modest per-period participation cost can explain the bulk of non-participation of US households. Indeed, when conducting the back of the envelope calculation of Vissing-Jorgensen (2002) with our Swedish cross-section we find that costs of 528 USD.[1] can explain 75% of non-participation. The theory used to generate this conclusion suggests that a large, positive wealth shock should greatly increase subsequent participation. Given that we our reduced form analyses don't support this prediction, it is natural to estimate how large

[1] The model is estimated and costs calculated in 2010 cpi adjusted Swedish Krona. Here and for the remainder of this paper, we use the Dec. 31, 2010 exchange rate of 6.72 SEK/1 USD when providing results in USDs.

the costs must be to generate the estimated responses. We find that necessary costs to generate non-participation are far larger than any previously estimated, with a median cost of entry for non-participants of almost 500,000 USD. Our basic conclusion – that substantial entry costs are required to account for non-participation in a model identified using the exogenous wealth variation – starkly contradicts the findings of most previous studies (e.g., Vissing-Jorgensen (2002); Gomes and Michaelides (2005); Alan (2006); Khorunzhina (2013)).

Evidence that participation responses to exogenous wealth shocks costs are too low to be explained by fixed participation costs was previously found in Andersen and Nielsen (2011). We build upon this study in several ways. First, given the high confidence in the causal interpretation of our reduced form estimates, our study validates their main conclusion that wealth has a small effect on stock market participation. More importantly however, the structural estimates in this study demonstrate how large fixed costs of entry and participation must be to rationalize observed participation responses. This allows us to reject the theory that financial costs are financial in nature despite frequently being modeled as such.

As an alternative explanation of non-participation, we re-estimate the model allowing for heterogeneity in the perceived equity premium. This exercise finds that beliefs of lower excess returns than what are observed historically are capable of generating the participation responses observed in our study without resorting to absurd costs of participation. Furthermore, the distribution of beliefs that we estimate is remarkably similar to the stated beliefs of equity returns documented in Hurd, Van Rooij, and Winter (2011). While this is a simple reinterpretation of the disincentives to participation, the consistency of the estimated belief distribution with other sources of stated beliefs suggests that informational frictions and lack of financial literacy, as suggested in Van Rooij, Lusardi, and Alessie (2011) and Grinblatt, Keloharju, and Linnainmaa (2011a)) are more believable causes of non-participation.

The remainder of the paper is structured as follows. Section 2 describes construction of lottery players by matching administrative data on participants in three lotteries to Statistics Sweden's register data on wealth. In describing our lottery samples, we address several important issues about external validity that are often raised about studies of lottery players. Importantly, Section 2.2 lays out our basic identification strategy, and Section 2.5 relates our strategy to previous calculations of fixed costs. In Section 3 we report the results from our reduced form analyses. Section 4 presents a structural model of life-cycle asset market participation and estimates the distribution of entry and participation costs implied by our variation in wealth. Section 4.4 extends this model to allow for heterogeneity in the cost distribution and uses reduced form heterogeneity analyses to identify heterogeneity in costs. Section 5 examines alternative explanations of non-participation and estimates the distribution of subjective beliefs of the equity premium that is consistent with the estimated effects of wealth, while Section 6 concludes.

2 Data and Identification Strategy

Our analyses are based on three samples of lottery players who have been matched, using personal identification numbers (PINs) to administrative records covering the entire Swedish population. Below, we begin brief description of the register variables that play a key role in our analyses.

We draw primarily on high-quality information about year-end financial portfolios (assets and debt) which are available 1999–2007. Until 2007, household wealth was taxable under Swedish tax law. To implement this tax, Statistics Sweden collected information from other branches of government, as well as banks and other financial institutions. A register known as the *Swedish Wealth Registry* contains detailed information about the year-end financial portfolios of the entire Swedish population during the study period. The register contains individual-level variables measuring bank account balances, mutual funds, directly held stocks, bonds, money market funds, debt, residential and commercial real estate, and other financial and real assets. This permits construction of several outcome variables that will use for the remainder of this paper, including our measures of stock market participation. Although originally, collected for tax purposes, the high quality of the records has made them extremely useful resource for researchers, as exemplified by several recent studies such as Calvet, Campbell, and Sodini (2007, 2009) and Calvet and Sodini (2014).

Several of other analyses also make use of the a rich set of demographic covariates available in Statistics Sweden's *Integrated Database for Labour Market Research*, which contains annual information (1990-2010) about a number of demographic characteristics which include income, employment, educational attainment, region of residence, retirement status and household composition.

Deciding whether the appropriate unit of study is a household or an individual depends on the extent to which it is reasonable to assume that the adults in a household make joint financial decisions. In our data, the wealth of a winning player's spouse or partner increases by about 20% of the total prize amount in the year of the win, often because the prize money won is deposited into a joint account. In the main analyses that follow, we therefore make the household the unit of analysis. A household always comprises one or two adults. Following Statistics Sweden, we say that two adults form a household if they are either married or cohabiting with an individual with whom they have at are either married or cohabiting with an individual with whom they have at least one child. All other adults are treated as one-person households. In Appendix C we show that our main conclusions are substantively identical if we exclude asset ownership through spouses from the study.

2.1 Lottery Data

We next turn to a description of the lottery data. Our basic 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. This sample includes three distinct samples of lottery players. The first is a monthly Swedish subscription lottery called Kombilotteriet ("Kombi"). Our second sample, Triss, contains of scratch lottery players who qualified for a TV show where they could win substantial amounts of money. Our final sample is a panel of individuals with prize-linked savings (PLS) accounts. PLS accounts are savings accounts which, instead of just paying interest, also incorporate a lottery element by enrolling account holders in lotteries. We describe all three samples briefly below, and refer the reader to Cesarini, Lindqvist, Östling, and Wallace (2013) for a richer description.

Kombi

Kombi is a monthly subscription lottery whose proceeds are given to the Swedish Social Democratic Party, by far the most dominant political force in Sweden during the post-war era. Subscribers choose their desired number of subscription tickets and are billed monthly usually by direct debit. Kombi provided an unbalanced panel covering our entire sample period. For each draw, the panel contains has one entry per eligible participant, and lists the players' PIN, number of tickets purchased and the prize amount won (for all prizes exceeding 1M SEK, net of taxes). For a small number of individuals ($\sim 1\%$) the PIN is missing and we do not include these individuals when constructing our final estimation sample.[2]

In each draw, every purchased ticket is assigned a unique number by Kombi, and the winning tickets are then drawn randomly from the set of purchased tickets. Therefore, two individuals (or households) who purchased the same number of tickets in a given draw face the exact same probability of winning a large prize. Because our main analyses are conducted at the household

[2]Because missingness is determined entirely by whether the participant supplies a valid PIN at enrollment, this restriction does not introduce any sample selection biases that would jeopardize the interpretation of our parameter estimates as causal. The restriction does change the composition of the sample for which we are estimating the treatment effect.

level, our empirical strategy is to compare each household winning a large prize with "matched control households" who did not win a large prize but who purchased exactly the same number of tickets in the month of the draw. To construct the cells, we began by computing the number of tickets owned by the household of each winning player. We then matched each large-prize winner to (up to) 100 non-winning households who did not win a large prize in the month of the draw.[3] The Kombi sample contains 46,486 total prizes including 339 large prizes.

Triss Sample

Our second sample is called Triss, a scratch-ticket lottery run since 1986 by Svenska Spel, the Swedish government-owned gambling company. Participants can win the opportunity to participate in a TV show (TV-Triss) where they can win a substantial prizes. Each month, around 25 TV-Triss prizes are awarded on television.

At the show, participant draws a prize from a stack of tickets. This stack is determined by a public prize plan that is subject to occasional revision. Because the tickets in the stack are shuffled and look identical, the prize won by the participant in the show is random conditional on the prize plan. Tv-Triss Prizes are paid out as a lump-sum and vary in size from 50,000 SEK to 5 million SEK (net of taxes).

Svenska Spel supplied us with information about all individuals who participated in the TV show between 1994 and 2010. With the help of Statistics Sweden, we were able to to use the information in the spreadsheet (name, age, region of residence, and often also the names of close relatives), to reliably identify the PINs of 98.7% of show participants. In the Online Appendix, we provide a detailed account of the processing of the data. The spreadsheet also notes any instances where the participant shared ownership of the ticket. Our analyses below are based exclusively on participants who did not indicate that they shared ownership of the winning ticket, but results do not change appreciably with these individuals included.

Our empirical strategy makes use of the fact that, conditional on the prize plan and winning exactly one prize, the nominal prize won is independent of pre-determined characteristics. To account for small changes in the real value of the prizes induced by inflation, we further restrict our comparison to individuals who won in the same year. Thus, our empirical strategy is to exploit the prize variation between individuals who won in the same "draw", where we define each unique combination of year and prize plan as a separate draw. If the members of a household win more than one prize in any given draw. In principle, households winning two prizes could be compare to other households who won two prizes in the same draw, but multiple wins are so rare that it is never possible to identify a successful match. After dropping prizes won by individuals whose PIN could not be reliably identified and players whose tickets were jointly owned, the final sample contains 4250 households.

PLS Sample

A PLS account is a savings account whose owner is enrolled in regular lotteries with monetary prizes (Kearney, Tufano, Guryan, and Hurst (2010)) paid in addition to (sometimes in lieu of) interest payments. Such accounts have existed in Sweden since the 1950s (Regeringen, 1972). The subsidies ceased in 1985, at which point the government authorized banks to offer prize-linked-savings products. Two systems were put into place. The savings banks (Sparbankerna) started offering their clients PLS-products through a system known as the Million Accounts ("Miljonkontot"), whereas the remaining banks joined forces and offered a PLS product known as Winner Accounts ("Vinnarkontot"). Approximately one in two Swedes held a PLS account.

During the period we study, 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 account

[3] The exact matching procedure is described in the Online Appendix.

balance (an account holder got one lottery ticket per 100 SEK in the account). Fixed prizes were prizes whose magnitude was not determined by the account balance of the winning account. The size of the odds prizes, on the other hand, depended on account balance. In each draw, an account's balance was proportional to the number of lottery tickets assigned to the account. An overwhelming fraction of the prizes awarded were small fixed prizes (typically 1000 SEK), but about XX% of the total sum of prizes awarded came from large prizes (100K SEK or more).

Our final lottery sample is obtained by combining data from two sources of information about the Winner Accounts: a set of printed lists with information about all prizes won 1979–2003 and microfiche images with information (account number, account owner's PIN, and number of tickets received) on all accounts in the draws between December 1986 and December 1994 (the "fiche period"). Both sources were retrieved from the Swedish National Archives.

Our final estimation sample contains both fixed and odds prize winners, although only odds prizes were awarded after XXX. The sample was constructed in two steps. In a first step, we digitized all the information on the prize lists. For each draw, these list all winning accounts (account number) and prize(s) won (type of prize, prize amount). The fiches list the account number and the PIN of each account owner, while the prize lists only list account numbers. It is therefore only possible to map each account number of prize-winning account during our sample period to a PIN if the account was open for at least part of the fiche period. In a second step, we therefore dropped prizes won by individuals whose PINs could not be identified.

[UPDATE THIS PARAGRAPH TO INCLUDE ODDS PRIZES]. Our identification strategy instead exploits the fact that in the population of households who won exactly n fixed prizes in a particular draw, the total sum of fixed prizes won is independent of account balances (and all other predetermined characteristics, including the date on which the account was opened). For each draw, we therefore assign two winning households to the same cell if they won an identical number of fixed prizes in that draw. This strategy is similar to that used by Imbens, Rubin, and Sacerdote (2001), Hankins, Hoestra, and Skiba (2011), and Hankins and Hoestra (2011) but unfortunately is only valid for fixed prizes. In a third step, we therefore drop odds-prizes from the sample period. In the Online Appendix, we show why our identifying assumption is valid for the subsample of fixed-prize-winning accounts that were in existence during the fiche period. This study uses a total of 311,331 PLS prizes (1999-2003), including 339 large prizes.

2.2 Identification Strategy

Our identification strategy thus uses the available data and knowledge about the institutional details of each lottery to define subsamples/cells within which wealth is assigned independently of potential outcomes. Table 1 summarizes the previous section's discussion of how these cells are constructed in each of the three lotteries. Normalizing the time of the lottery to $s = 0$, our main estimating equation is given by,

$$Y_{i,s} = \beta_{1,s} \times L_{i,0} + \mathbf{Z}_{i,-1} \times \gamma_s + \mathbf{X}_i \times M_s + \eta_{i,s} \tag{1}$$

where i indexes households, $L_{i,0}$ denotes the prize amount won (in 2010 SEK), X_i is a vector of cell fixed effects and Z_i is a vector controls observed the year before the lottery. The key identifying assumption needed for β_1 to have a causal interpretation is that the prize amount won is independent of $\eta_{i,s}$ conditional on the cell fixed effects. In practice, we often control for several additional characteristics ($\mathbf{Z}_{i,-1}$) which are always measured in the year before the lottery. These controls are included to absorb more variance of the residual and hence improve the precision of our estimates. We estimate separate equations for participation in the year of the lottery ($s = 0$) and various s year horizons. Since wealth data is only available starting in 1999, analysis in which $\mathbf{Z}_{i,-1}$ includes financial variables requires restricting the sample to post-1998 winners.

To get a better sense of the source of our identifying variation, Table 2 provides basic information about the distribution of prizes won by the players in our samples. For each lottery, and the pooled sample, the table shows the total number of small (<100,000 SEK), medium (<1,000,000 SEK) and large prizes (>1,000,000 SEK). To put these prize amounts in perspective, the median annual after-tax earnings of a Swede working full time was roughly 170,000 SEK, in 2010 prices, in 1998 (the first year of the sample period). A first important message from Table 2 is that even though the number of prizes won vary dramatically across the lotteries, all lotteries contribute substantial identifying variation to our study. The total value of the after-tax prize money disbursed to the winners in our samples is almost 3.5 billion SEK (about 500 million dollars).

A second important message is that effects we report in the paper therefore assign relatively little weight to the marginal effects of small lottery prizes, even though these account for a large fraction of the number of prizes won. The reason is that even though a large number of prizes are small, they account for only a modest fraction of identifying variation. For example, dropping all the prizes below 10,000 SEK from the sample reduces the total amount of treatment variation by 19%. In Kombi, all of the identifying variation comes from comparisons of players who win large prizes to players who did not win a prize. In Triss, most identifying variation comes from comparisons of winners of large prizes to winners of small or modest prizes. Finally, in PLS, virtually all of the identifying variation comes from winners of medium or large prizes to winners of small prizes. Consequently, our estimates are most informative about the impacts of wealth shocks equal to several years of income, for example the effect that major changes to capital income taxes or pension systems can have on lifetime wealth.

2.3 Internal Validity

To test our key identifying assumption, we ran quasi-randomization tests premised on the simple idea that if lottery wealth is random conditional on the cell fixed effects, it should not be possible to predict the lottery outcome using covariates determined before the lottery in a regression that controls for the cell fixed effects. We estimate the following regression equation,

$$L_{i,0} = X_{i,0} \times \Gamma + Y_{i,-1} \times \rho_{-1} + \epsilon_i \tag{2}$$

where X_i denotes the individual's assigned cell, $Y_{i,-1}$ is a set of time-invariant characteristics (sex and birth year), as well time-varying characteristics measured in the year before the lottery. These lagged characteristics include marital status, educational attainment, income and a host of financial characteristics. As shown in Table **??**, none of the predetermined characteristics are significant predictors of the prize amount, individually or jointly, once the cell fixed effects are included as controls. This result holds across all three lotteries, the pooled sample, and the sample of post-1998 winners.

2.4 External Validity

One concern frequently voiced about studies of lottery players is that individuals who play the lottery may not be representative of the population. To investigate the representativeness of our samples, we compare our samples to a population sample of adult Swedes matched on age and gender from the Swedish population the year of the win. The results are presented in columns 1 and 2 of Table 4. We observe very few differences amongst demographic variables except that winners show a slightly higher tendency to be born in Sweden than the general population. It is natural to suspect that this reflects a lower tendency for immigrants to participate in lotteries, although this can't be confirmed. In addition, we observe that our lottery sample has a slgihtly higher income than the general population.

To avoid comparisons of outcomes that may be endogenous to the outcome of the lottery, when analyzing wealth measures we restrict our sample to individuals that won later than 1999. As Table 4 shows, there are again very little differences in demographic variables other than those that are observed in the full sample. When comparing wealth variables, we observe that winners tend to be slightly wealthier, have less debt, and slightly more likely to participate in the stock market than the matched population, although these differences tend to be small. Most importantly however, these differences disappear when the PLS sample is excluded. This is unsurprising given that this group self-selected based largely upon having a bank account and therefore positive financial wealth.

To probe further into our sample's representativeness, we estimate a cross-sectional probit regression of stock market participation (defined as owning stock or mutual funds) on household characteristics to see if the patterns in our sample of lottery players resembles those in representative sample. Specifically, we replicate columns 1-3 of Table 6 in Calvet, Campbell, and Sodini (2007) for our sample of post-1998 lottery players and matched sample and compare to the results of this influential paper. We again restrict our samples to winners after the first year we have reliable wealth records to avoid using wealth variation that was induced by the lottery (and may therefore change the coefficient estimates even if the a lottery population were representative).

Table 5 presents this replication, with columns 1 presenting our estimates, column 2 presenting the corresponding z-statistic, and 3 presenting the effect of increasing each variable 1 standard deviation from its median level or changing an indicator from zero to one. Columns 4-6 and columns 7-9 present the same estimation for the matched population sample and that of Calvet, Campbell, and Sodini (2007) respecitively. Generally, the estimated effects are quite similar between our lottery sample and Calvet, Campbell, and Sodini (2007) with three exceptions. First, we omit a variable relating to private pensions and income due to not understanding how this was constructed.[4]. Second, we observe that being an immigrant is associate with significantly lower probability of stock ownership in the population, but has little effect in our lottery sample. This likely reflects the previously documented lower prevalence of immigrants in our sample of lottery winners. Finally, we observe different effects of missing education. We suspect that his relates to differences in our coding of this variable. Absent these three effects, we observe little difference in the cross-sectional relationship of our lottery sample and the population. In comparing our lottery sample to our matched population sample, we again observe similar estimated effects. Thus, the results in Table 5 give no strong reason to suspect that our sample of lottery winners differs substantially from the population outside of the manners previously documented.

2.5 Generalizing Beyond Sweden

Finally, an important concern about external validity is that insights from Sweden may not generalize to other countries. There is surely some merit to the view, which is also discussed by Calvet, Campbell, and Sodini (2007) (p. 712). We nevertheless believe there are compelling reasons to expect the findings we report here to be relevant beyond the Swedish setting. For example, previous work has noted that the predictors of non-participation in Sweden are surprisingly similar the United States (Calvet, Campbell, and Sodini (2007), Table I). Cross-sectional analyses also find that the composition of Swedish household wealth is no outlier when compared to what has been observed for other industrial countries. For example, the fraction of non-participation households was 62% in Sweden in 1999, compared to 59% in the US the same year. These similarities suggest to us that it is plausible to expect that the causal processes that give rise to non-participation in the two countries are not fundamentally different.

To provide some more indirect evidence on generalize-ability, and set the stage for the rest of the paper, we now give a simple illustration of how researchers have sought to improve our understanding of non-participation by augmenting the standard household finance model with participation costs. We also show that when the models are calibrated to match the Swedish distribution of wealth and degree of non-participation, the required costs are of similar magnitude to those reported for the US by Vissing-Jorgensen (2002).

Vissing-Jorgensen (2002)'s presents an influential calibration exercise that demonstrates why

[4]Future versions of this paper will hopefully include this variable

low costs of participation might be assumed from the cross-sectional wealth distribution. Assuming time separable and homothetic preferences, household will not participate in equity markets if the per period cost of doing so is greater than the expected gain. Defining the certainty equivalent of return r_{ce} as the certain rate of return that makes an individual between investing in risky assets and not participating, a household's per period benefit from participating in equity markets can then be approximated as

$$Benefit_{i,t} = W_{i,t} \times \alpha_i \times (r_{ce,t} - r_f) \tag{3}$$

where $W_{i,t}$ denotes individual i's wealth, α_i denotes individual i's portfolio allocation, and r_f denotes the risk free rate. Vissing-Jorgensen (2002) assumes $r_{ce} = .05$, $r_f = .01$, and $\alpha_{i,1993} = .566$ is calibrated to the median of risky asset share of financial portfolios in the 1993 PSID. Benefits are then a linear function of wealth, and given that the 75^{th} percentile of wealth amongst non-participants is approximately $14,900$ USD (adjusted to 2010 price levels) in the 1993 PSID, this implies that an annual cost of $F_i^P = \$340$ is sufficient to explain non-participation for 75% of non-participants.

It is straightforward to repeat this calculation in our 1999 cross-sectional data of lottery players. Assuming again that $r_{ce} - r_f = .04$ and calibrating $\alpha_i = .43$ to match the median cross-sectional equity share of household financial wealth in our sample, Figure 1 presents the cross-sectional distribution of particiaption costs necessary to rationalize non-participation. Here we observe that for the 75^{th} percentile of our wealth distribution the necessary per period cost of participation is 3351 SEK (528 USD), not dramatically different from the 340 USD estimated in Vissing-Jorgensen (2002). Appendix A presents a second calibration of the participation cost CDF implied by the lottery sample's 1999 cross-section, and again finds it to be quite similar to that calculated in Vissing-Jorgensen (2002).

These exercises demonstrate two things. First, simple analysis of cross-sectional patterns in our lottery sample is consistent with the same patterns documented in the United States. More importantly however, using only cross-sectional variation in wealth amongst pre-win lottery players would lead us to conclude that moderate fixed costs are capable of explaining non-participation in equity markets. We will revisit these cross-sectional cost estimates in section 4 to demonstrate their implications for participation responses for our sample of lottery winners. Except for this exercise, the remainder of this paper we will use within individual variation of wealth to estimate participation responses and costs of entry and participation. We will show that this variation in wealth results in higher estimated costs of participation than have been previously found in the literature, and as a result bring into question the structural interpretation of previous cost estimates based on cross-sectional variation.

3 Empirical Results

In this section we use equation 1 to estimate the causal impact of wealth on stock market. In all results presented in this section, we normalize winnings by 1M SEK so that the estimated $\beta_{1,s}$ coefficient can be interpreted as the effect receiving 1M SEK (approximately 150,000 USD) has on the probability of participation in stock market participation. More specifically, this coefficient can be interpreted as the change in probability of participation for an individual that receives 1M SEK relative to an identical individual that received nothing.

In our primary specification, we estimate the unconditional effect of wealth on eqauity market particiaption if participation is an indicator variable equal to 1 if the household's year-end portfolio included any directly or indirectly held stocks (and 0 otherwise). We control for cell fixed effects and a handful of predetermined characteristics measured at $s = -1$. Column one of Table 6 presents the full regression for $s = 0$, finding an estimated coefficient of .038. This suggests that (ignoring aggregate effects) that providing everyone with 1M SEK would increase participation in equity markets by a fraction of .038 the year of wealth assignment. This effect, although quite small, is significant at a 1% level. In column five of table 6, we estimate this same specification except with a more restrictive definition of equity market participation of direct stock ownership. With this participation definition, we estimate that receiving 1M SEK causes an increase in participation probability of .020 the year of win, although it is not statistically significant.

Figure 2 depicts the estimated causal effect of wealth from a linear probability model estimated using our pooled lottery sample for up to 10 years (at $s = 0, ..., 10$). Panel A presents the estimated effect when equity market participation includes both directly or indirectly held stocks. The estimated effect appears to be immediate and permanent, as the coefficient varies between .045 and .052 while remaining significant at all horizons. If If we instead adopt an event study framework, and impose the restriction that $\beta_{1,s} = \beta$ for all $s = 0, ..., 10$ the estimate is .043 with a standard error of .007. It is useful to benchmark this estimate against the cross-sectional relationship between wealth and participation. In our data, the cross-sectional effect of a one standard deviation increase in wealth on participation probability is .219, while the causal event study estimate implies an increase in participation probability of less than .04 due to the same increase in wealth. Panel B presents the estimated effect when the definition of equity market participation is restricted only to directly held stocks. We observe some evidence of a delayed effect, as the estimated effect on participation only becomes significant in the year following the win. Although the effect of receiving 1M SEK remains significant at all subsequent horizons, it remains small and varies between .02 and .045. These estimates suggest that while the effect on direct stock market participation may be partially delayed for one year, it is otherwise permanent, significant, and small.

The small aggregate effect may mask substantial heterogeneity, and now explore potential differential responses in greater detail. Given the well documented inertia in participation (see **?**), we began by investigating how responses vary when the sample is stratified by stock participation in $s = -1$. The results if participation is defined as direct or indirect stock ownership are presented in column three of Table 6. For those individuals that were not participants before wealth assignment, the estimated impact is an increase in participation probability of .113 (s.e.=.02) in the year of the lottery. In the population of participants, the estimated response is positive but very small and not significant. As shown in Figure 9, these patterns remain virtually unchanged up to four years following assignment of wealth.[5] Thus, virtually the entire participation response is driven by nonparticipants, a finding consistent with the predictions of a model in which, large, one-time, fixed costs of entry feature prominently. Column six of Table 6 presents the same estimation for direct stock market participation, and shows that the estimated effect is not significant for either prior participants or non-participants.

Given the evidence that the treatment response is explained almost entirely by a positive effect of wealth in the population of nonrespondents , all tests for heterogeneous treatment effects in the analyses that follow are based on nonparticipants. In addition, because there do not appear to be economically or statistically significant impacts on direct stock ownership, all subsequent analysis uses the more expansive definition of stock ownership. Finally, because the effect appears to be one time and permanent, we only present results for the year following win. Effects for non-participants as well as other horizons are presented in Appendices C.2 and C.

We test for heterogeneous responses in subsamples stratified along various characteristics, including home ownership, debt, self employment, recent stock market performence, gender, age, and educational attainment. In each heterogeneity analysis, we are conceptually interested in comparing the estimated effect of wealth in subsamples stratified along one of the dimensions, for example winners with and without a college degree. Procedurally, we run a single regression in which all regressors are interacted with indicator variable(s) for the subpopulations. The pooled regression recovers exactly the same coefficient estimates as those obtained when Equation 1 is estimated separately in each of the subsamples. To test for heterogeneity, we conduct an F-test of the null

[5]Note that because we are conditioning on participation status in 1999, in this figure the sample size decreases with horizon. This accounts for the observed increase in standard errors over time.

hypothesis that the coefficients are identical.

Because only wealth is randomly assigned, evidence of treatment effect heterogeneity along some dimension X need not imply that varying X exogenously will change participation costs. For example, treatment effect heterogeneity by college attainment could in principle arise because college completion is correlated with some factor (pre-college ability) that reduces participation costs independently of college completion. In addition, because participation status is not randmonly assigned and the composition of participants and non-participants in each subsample differs, treatment effect heterogeneity could reflect individual heterogeneity that caused selection into the appropriate participation status. Nevertheless, the heterogeneity analyses nevertheless provide useful information about how participation costs are distributed across observable characteristics. Such information is valuable for formulating new hypotheses about the sources of heterogeneity in participation costs, and is a key input into to our treatment of cost heterogeneity in the structural model.

Table 7 presents the estimated causal effect of wealth on participation status for pre-lottery non participants from these heterogeneity analyses. We begin with examining how home ownership impacts the effect on participation probability. A priori unclear whether home ownership should cause stronger or weaker increases in participation probability. On one hand, to the extent which real estate is a risky investment, home owners are already exposed risk and thus might be less inclined to take on equity risk. This effect has previously highlighted in Grossman and Laroque (1990) , Flavin and Yamashita (2002), and Cocco (2005). In addition, such individuals may rationally choose to use assigned wealth to upgrade their home. On the other hand, non-home owners that are non-participants in equity may choose to forego equity market participation and instead purchase a home. Estimates suggest that pre-lottery home owners exhibit an increase in equity market participation probability of .104, while non-home owners exhibit an increase of .142. Thus, the treatment effect is higher for non-home owners, although the difference is not significant. We will examine housing effects in greater detail in section 5, but this estimation suggests that heterogeneity in the effect of wealth on participation between home owners and non-homeowners is limited.

We next examine how the effect of wealth differs for individuals with and without pre-existing debt. The rate at which individuals can borrow typically exceeds the risk free rate which, as shown in Davis, Kubler, and Willen (2006), creates a "borrowing wedge"that may explain why indebted households elect to repay debt our participating in equity markets. Column 1 shows that the estimated effect of 1M SEK on the participation probability is 0.211 in households classified as debt-free (<10K SEK in debt), compared to 0.093 in all other households. This difference is statistically significant. Our next heterogeneity analyses is inspired by research suggesting that individuals may rationally choose to not participate because of uninsurable labor income risks Viceira (2001) Heaton and Lucas (2000). We therefore estimate the effect in nonparticipants that are and are not self-employed. As shown in Column 3 and 4, we find that no evidence of an increase in participation probabilities in households where the winner was self-employed, and a strong response in all other households.

Column 5 shows the results in subsamples stratified by recent stock market performance, distinguishing players by whether they won following a year in which market returns were positive or negative. If individuals' subjective beliefs about the expected equity premium are influenced by recent aggregate returns, then the participation response may too. We find that the effect of wealth on participation is indeed weaker amongst lottery players who won following a year of poor equity performance. Such winners exhibit only a .055 increase in equity market participation, while winners following a year of positive equity returns exhibit a .140 increase in participation due to receipt of 1M SEK.

In the bottom row of Table 7 we report heterogeneity analyses for the remaining characteristics. Columns 9 and 10 show that the impact of 1M SEK on the participation probability of males is .147 as opposed to .112 for females. Although Barber and Odean (2001) have previously documented gender differences in portfolio decisions, we find no significant differences in the treatment effects. Column 11 and 12 show that younger individuals are more affected than older individuals, although the effect is not statistically significantly different. These coefficients are consistent with a one time fixed entry cost that, all else equal, would make older workers less likely to enter given they have fewer remaining years in which to harvest the gains of participation. Finally, columns 13 and 14 show that individuals with a college degree exhibit an increase in participation probability of .230 compared to .09 in individuals without a college degree. This difference is significant, and consistent with theories that intelligence and cognitive constraints impact participation (Grinblatt, Keloharju, and Linnainmaa (2011b), Van Rooij, Lusardi, and Alessie (2012)), assuming that attending college relaxes cognitive constraints.

The heterogeneity analyses thus finds that the effect of wealth on stock market participation generally varies in intuitive, meaningful ways that are consistent with previous theories of stock market non-participation. In section 4.4 we will revisit these effects and estimate what they imply for different costs of participation and entry.

3.1 Are the Effects Nonlinear?

Under our identifying assumption, our estimator gives en unbiased estimate of a weighted treatment effect, but the linear estimator will assign most weight to the marginal effect of wealth at modest to large wealth shocks, as such prizes account for most of our identifying variation. One possible interpretation of the discrepancy between our causal estimates and the cross-sectional estimates is that the effects of wealth may be non-linear. Additionally, models with participation costs of the type discussed in Vissing-Jorgensen (2002) have the additional property that households should follow a threshold strategy: for each household there exists some wealth level above which participation is always optimal. Such a threshold rule is likely to show up in the form of non-linear effects.

To test for non-linear effects, we modify our basic estimating equation so that it can accommodate non-linear responses in a fairly transparent and easy-to-interpret way. Specifically, we replace the continuous prize variable by indicator variables for the prize amount won. By estimating the effect of wealth at different thresholds, it is possible to identify non-linear effects. Column 4 of Table 6 presents results from the regression with thresholds at 100k SEK, 1M SEK and 2M SEK. A prize of 100k-1M SEK increases the participation probability by 0.060, a prize of 1M-2M SEK increases participation probability by .185, and a prize of more than 2M SEK increases participation probability by .325. Normalizing these estimates by the median prize size in each group (respectively 110k SEK, 100k, 1.1M SEK and 2.7M SEK) results in linear equivalent coefficients of .545, .163, and .120. This suggests that the marginal effect of wealth on participation is is rapidly diminishing amongst pre-lottery non-participants.

3.2 Calibrating Participation Costs to Match Causal Estimates

The causal effects we estimate in non-respondents are not easy to reconcile with the hypothesis that most of non-participation is due to households facing a modest ongoing participation cost (of the order a few hundred dollars per year). Under this theory, many households decline to participate for the simple reason that at their level of wealth, the gains from participation (which are proportional to wealth) do not offset the costs of.04*.43*1.1 participation (which are fixed). But as we now show, our estimated participation responses to the wealth shocks are far smaller than is predicted by a model with an annual cost of a few hundred dollars.

To illustrate, we use the effects we estimate in our sample of nonparticipants. We conservatively assume that all households have zero wealth prior to the lottery. Evaluating Equation 3 at the median wealth levels for winners of 1-2M SEK and 2M+ SEK, we have that the per-period benefit of participation is 18,920 SEK (2823 USD) at a wealth level of 1M and 46440 SEK (6930 USD), respectively. Given that the increase in participation for each of these groups is .185 and .325 respectively, this would suggest that per-period participation costs of at least 2823 USD are required to rationalize 81.5% of non-participation, and costs of at least 6930 USD are required to rationalize 67.5% of non-participation. In comparison, to the costs that were calculated in section 2.5 as necessary to explain 75% of non participation in the US and pre-lottery cross-sections 340 USD and 528 USD. Thus, the necessary per period costs to explain non-participation effects estimated from wealth due to lottery variation are an order of magnitude larger.

3.3 Robustness Checks

We conducted a number of sensitivity checks to explore the robustness of our results to household definition, definition of the dependent variable, sample selection criteria, and choice of estimator. We report the results from these robustness analyses in Appendix C. For each of these variations, we report the causal coefficient estimate of the uncondtional effect and of the effect conditional on participation and non-participation at the $s = 0$ and $s = 3$ horizons. We find that restricting household definition to exclude spouses slightly lowers effect estimates, but changes no conclusions from our baseline estimation. We find that using a restrictive definition of participation results in much lower and generally non-significant coefficient estimates, suggesting that households participate primarily by purchasing mutual funds. We find that while the estimated effects vary some across lotteries (notably, the estimated effect on participation is larger for the Kombi sample), the results are consistent with those of the pooled sample. Thus, our main conclusions are not driven by any single lottery. Finally, we find that results from the Probit analogue, and their standard errors, are very similar to the linear probability model estimates reported in our primary specification. In short, our main findings appear robust to a variety of alternative specifications and analyses.

3.4 Summary of Reduced Form Findings

To summarize, our reduced form analyses shows that the effect of wealth on participation is immediate, permanent, heterogeneous, statistically significant, but small.

The effect is *immediate*, as it is usually discernible in the year of the lottery and it appears *permanent*, as the estimated effects of wealth on participation in the years following the lottery are of similar magnitude.

The effects are *small* both when bench-marked both against cross-sectional estimates and the participation responses predicted by a standard household finance model augmented with modest participation costs. The participation costs that would be required to explain the participation responses we estimate are thus strikingly high compared to earlier work.

The effects are *heterogeneous*. The heterogeneous effects we observe are intuitive and usually easy to reconcile with standard theories of participation. Most importantly, we find that the increase in participation probability of .038 per million SEK in our pooled sample masks substantial heterogeneity. In the linear model, the effect of 1M SEK is negligibly small in households who already participated, whereas it increases the probability of participation in the approximately 35% nonparticipating households by 12 percentage points. We also find a more elastic response in individuals in college-educated winners, in winners with lower background risk, and in households with lower debt and who win during an economic expansion.

4 A Structural Model

In this section we present a structural model of equity market participation and use it to estimate the financial costs of participation consistent with our reduced form findings. The structural model has several aspects that make it especially attractive. First, in our estimation of the effect of wealth assignment we distinguish between wealth that is assigned and pre-existing wealth, finding that both are associated with higher participation. However, most economic theories, including our structural model, treat total wealth as the key state variable in participation decision. Second, given the estimated non-linear effects, it is not clear escatly how wealth affects consumer decisions. Modeling the participation decision rule takes this relationship into account. Finally, the estimates presented in previous sections do not map cleanly to costs. In particular, absent a modeling framework, it is impossible to distinguish between entry and participation costs. The structural model allows for and permits estimation of both costs of entry and participation.

After presenting the model, we will undertake a series of exercises. First we calibrate it to the cross-sectional implied costs of participation calculated in Section 2.5 and simulate random assignment of wealth, including each household's participation decision after receipt of wealth. We then repeat estimations presented in Section 3 on our simulated data set, and find that the effect of wealth on participation probability estimated from the model are significantly higher than those estimated in the data. In a second exercise, we estimate costs of participation and entry that best replicate the participation responses of lottery winners using the method of indirect inference. Intuitively, we choose participation and entry costs that best replicate the estimates shown in section 3. Finally, we allow for heterogeity in the cost distribution according to the characteristics analyzed in Table 7 and use the corresponding estimated effects to identify heterogeneous costs of entry.

4.1 Model

We assume a very standard model of life-cycle saving, market participation, and portfolio choice. Each period an age t agent chooses how much to consume, save, and invest in equity markets according to their current resources. An agent has finite lifespan, living to age T, but faces mortality risk with exogenous survival probability from period t to $t+1$ denoted s_t. Upon death, an agent receives terminal payout of zero. If an agent decides to participate in equity markets they face separate financial costs of participation and entry. Participation costs, denoted κ, are payed each period an agent allocates non-zero wealth to equity holdings. Entry costs, denoted χ are paid whenever a non-participating agent decides to enter equity markets. Equity provides a return r_s with $\mathbb{E}(r_s) > r_f$, but the return is risky. In addition, each period an agent is endowed with age specific labor income y_t.

For an agent that decides to not-participate in equity markets, their value function (V_t^{NP}) is standard. Given a continuation value V_{t+1}^{NP}, the agent simply decides how much to consume and how much to save at risk free rate r_f.

$$\begin{aligned} V_t^{NP}(W_t) = & \max_{c_t, W_{t+1}} u(c_t) + \beta s_t \mathbb{E}_{y_{t+1}} \left[V_{t+1}(W_{t+1}, I_t) \right] \\ & W_{t+1} = r_f \left(W_t - c_t \right) + y_{t+1} \\ & I_t = 1 \end{aligned} \tag{4}$$

Here, I_t is an indicator of non-participation during year t. For an agent that participates in the equity market, the agent must pay costs of participation and decide how to allocate wealth between stocks and bonds. Given continuation value V_{t+1}, an equity market participants problem can be expressed as

$$\begin{aligned} V_t^{P}(W_t) = & \max_{c_t, W_{t+1}, \alpha_t} u(c_t) + \beta s_t \mathbb{E}_{y_{t+1}, r_{s,t+1}} \left[V_{t+1}(W_{t+1}, I_t) \right] \\ & W_{t+1} = r_f \left(W_t - c_t - \kappa \right) + \alpha_t (r_{s,t+1} - r_f) \left(W_t - c_t - \kappa \right) + y_{t+1} \\ & I_t = 0 \end{aligned} \tag{5}$$

Finally, each period an agent's decision to participate or not is determined by the max of the above two value functions. An agent's full decision problem is specified as

$$V_t(W_t, I_{t-1}) = \max\{V_t^{NP}(W_t)\ ,\ V_t^{P}(W_t - \chi I_{t-1})\}. \tag{6}$$

We assume that s_t, $r_{s,t}$ and y_t follow known stochastic processes. Although our data is collected at the household level, we treat s_t as the survival probability as a function of an individual at age t. This can be interpreted as either that the household dies upon death of the specified agent or that there is no preference for remaining household members upon death. In addition, we assume an age-dependent income process, with log income being expressed as

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

We assume that $f(t)$ is a quadratic in age and $\sigma_{y,t}$ is chosen to match the the age-specific dispersion in earnings. We omit any persistence in earnings and any heterogeneity in earnings or survival probabililties.Both of these processes are estimated from the 1999 cross-section of the Swedish population, with the exact procedure and resulting estimates presented in Appendix D.

We assume that the equity returns are lognormally distributed, with mean $\bar{r} = .065$ and standard deviation $\sigma_s = .21$. Furthermore, we calibrate $r_f = .01$ at the historical average of Swedish treasury bills. This parametrization is taken from **?**, which estimates the post-war equity returns of the Swedish stock exchange.

Finally, we assume that utility is CRRA, and calibrate risk aversion parameter $\nu = 4$. We could in principal estimate this, but choose to instead focus on estimating participation costs. We assume that entry and participation costs are distributed according to the following distributions

$$\begin{aligned} \ln \kappa_i &\sim F_{\theta_\kappa}(\kappa) \\ \ln \chi_i &\sim G_{\theta_\chi}(\chi) \end{aligned} \tag{7}$$

where θ_χ consists of parameters that characterize the population distribution of entry costs and θ_κ characterizes the population distribution of participation costs.

Subsequent analysis will focus on replicating the reduced form estimates preseted in section 3 on a data set simulated from our model. Having solved the model for the optimal consumption, savings, and participation decision rules, we simulate can simulate the model to generate a sequence of participation decisions in which we can replicate all reduced form analyses. This simulation process is standard except for the first year. During the first year each household's wealth is augmented with unexpected lottery winnings L_i. Thus, in the second period an agent solves the decision problem

$$V_t(W_t + L_i, I_{t-1}) = \max\{V_t^{NP}(W_t + L_i)\ ,\ V_t^{P}(W_t + L_i - \chi I_{t-1})\}, \tag{8}$$

although receipt of wealth L_i was not foreseen. In our baseline analyses we estimate the unconditional effect, effect conditional on prior participation status, and effect conditional on prize thresholds.

4.2 Calibration

As a first exercise, we use the results section 2.5 to calibrate costs of participation such that $\kappa =$3351 SEK (528 USD) and $\chi = 0$. In addition, in keeping with other studies in the literature we assume that individuals are not able to borrow against future labor income. Thus, the decision problem in equation 5 is augmented with the additional constraint that $\alpha \geq 0$.

Results from this exercise are presented in column 2 of Table 8. Here we see that in general the estimated effects are significantly larger than we observe in the data. The unconditional effect of 1M SEK on participation is more than double its empirical counterpart (.090 as opposed to

.038), as is the effect on non-participants (.261 as opposed to .057). In examining the effects of different prize sizes on the participation probability of non-participants, we again observe that the model predicts a far larger effect, especially amongst those that won prizes of 1-2M SEK. Finally, we while empirically we observe virtually no effect of wealth on continued participation, this model suggests positive effects. In the linear specification the model predicts an increase of .057 in participation probability amongst pre-lottery participants. For winners of prizes of 1-2M SEK (2M+ SEK), the model predicts an increase in participation probability of .175 (.155) relative to pre-lottery participants that received small prizes. This sharply contrasts empirical results which suggest a negligible effect on continued participation probabilities in all specifications. Finally, we note that for both pre-lottery participants and non-participants there is virtually no difference in the estimated effect of receiving 1-2M SEK and 2M+ SEK. This finding reflects that in this model the participation decision rule follows a threshold strategy in which individuals participate iff their wealth is above a certain threshold. In a calibration with a common fixed cost, there is no scope for differential effects of wealth above some threshold, resulting in common effects for recipients of 1-2M and 2M+ SEK.

This calibration provides strong evidence that a small fixed cost of participation is incapable of replicating the effects we estimate empirically. Although we calibrate the model according to Vissing-Jorgensen (2002), this calibration is very similar to other studies of life-cycle equity market participation. To highlight this, Table 9 presents a summary of costs of participation and entry that have been estimated in other studies. In all studies that use cross sectional variation in wealth to identify participation costs, the cost of participation (interpreted within our model) never rises above 1000 USD and the cost of entry never rises above 1300 USD. Thus, in all of these models and calibration one would expect similar effects of wealth that we find in this calibration. This finding is likely true whether or not housing is modeled given that studies that model housing have similarly small equity market participation costs. Furthermore, although we omit housing from the model, we will show in Section 5 that housing does not appear to be driving any of our results. Thus, the models previous studies will in general have a hard time replicating our empirical findings.

In this table, there are two studies that estimate substantially higher costs of entry and or participation. Khorunzhina (2013) uses the panel dimension to identify costs of participation and finds participation costs that are substantially higher than any other studies based on life cycle models. In particular, this study estimates a per-period participation cost of 4-6% of the permanent component of labor income, which varies between 1200 and 3050 USD in our model. In addition, Andersen and Nielsen (2011) use wealth variation most similar to ours and use back of the envelope calculations to calculate an implied per-period participation cost of 1550 USD or an implied entry cost of the annuitized value. The differences in conclusions of these studies relative to the others likely stems from the differences in the source of identifying wealth variation. While we suspect they would fare better at replicating our findings, in the coming section we will estimate exactly what are the costs of participation and entry implied by our study.

4.3 Baseline Structural Results

In our second and third exercises, we estimate the model using the method of indirect inference. For each individual we sample a from the cost distribution implied by parameters $\Theta = (\theta_\kappa, \theta_\chi)$, repeat the simulation and estimation procedure, and choose the parameter set Θ that most closely replicates the empirical estiamtes presented in column 1 of Table 8. Formally, the estimation procedure can be formally as:

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

where $\hat{\beta}$ represents the the vector of empirical estimates and $\tilde{\beta}(\Theta)$ represents the vector of coefficients implied by the model. Although W could be any positive semi-definite matrix, as a baseline we use the identity matrix. For further information on the implementation of our structural estimation procedure we refer the reader to appendix D.2.

We identify the population distribution of participation costs, as well as the distribution of entry costs non-parametrically. .To do so, we bound the set of feasible costs by assuming that $F_{\theta_\kappa}(\overline{\kappa}) = 1$ and that $G_{\theta_\chi}(\overline{\chi}) = 1$. Knowing further that $F_{\theta_\kappa}(0) = 0$ and $G_{\theta_\chi}(0) = 0$, we construct a linearly-spaced grid consisting of 7 points denoted as x^n_κ and x^n_χ respectively. Assuming that the CDFs are piecewise linear between these points and defining

$$\theta_\kappa = \{\theta^n_\kappa | F_{\theta_\kappa}(x^n_\kappa) = \theta^n_\kappa\}$$
$$\theta_\chi = \{\theta^n_\chi | G_{\theta_\chi}(x^n_\chi) = \theta^n_\chi\},$$

the cost distributions are 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.

It is clearly not possible to identify the population distribution of entry costs given that a significant proportion of our sample participate in the stock market prior to random assignment of wealth, and the estimated distribution should only be interpreted as reflecting the costs of pre-lottery non-participants. However, this identification challenge is not unique to our study, and is the primary reason why most studies only permit a single type of cost. We find that this specification is necessary however to match the participation responses of both pre-lottery participants and non-participants.

Column 3 of Table 8 and Figure 4 present the results of our baseline estimation. In Column 3 of Table 8 we see that the fit of our model is near exact. We reasonably match all coefficients, including the unconditional effect of wealth on equity market participation, the effect conditional on past participations tatus, near zero effect on continued participation, and nonlinear effects amongst non-participants. This demonstrates that the standard model is able to generate the empirical estimates with appropriately chosen costs.

The estimated distribution of these costs is presented in Figure 4. First, note that this estimation suggests that fixed costs of participation are indeed quite small, with a median estimated cost of 200 SEK (30 USD). Furthermore, 90% of individuals have participation costs less than 500 SEK (80 USD), and almost none have costs more than 1000 SEK (160 USD). However, the estimated entry cost distribution suggests an extreme importance o entry costs in the equity market participation. The estimated distribution suggests that 40% of pre-lottery non-participants have an entry cost of more than 670,000 SEK (100,000 USD), 50% have an entry cost of more than 2,700,000 SEK (400,000 USD), and a non-negligible proportion have an entry cost of more than 3,300,000 SEK (500,000 USD).

While these cost estimates may seem absurd, they should not come as a surprise. The reduced form estimates we target document little increase in participation following receipts of large amounts of money. Inducing non-participating households to continue to forego the significant equity premium given their new wealth necessitates a very strong disincentive to participate. Because wealth is estimated to have zero effect on continued participation and significant participation costs would cause a positive effect, per period particiaption costs must be small. This leaves entry costs as the primary instrument to deter entry into the market. The high costs implied by the estimated distribution reflect this deterrence. Thus, although it is difficult to imagine entry entailing a financial cost of 400,000 USD, the strong disincentive is necessary given the model and empirical patterns.

4.4 Structural Model with Heterogeneous Costs

As a third exercise, we augment our structure model to allow for heterogeneity in the distribution of costs amongst the various subsamples examined in Table 7. In this table, heterogeneity analyses are slightly difficult to interpret given that the underlying distribution of non-participants is correlated with the conditioning variables. Because the heterogeneous characteristics are not exogenously assigned, when we observe larger effect on participation probability it is unclear if the differential responses reflect different costs or differences in pre-assignment states. To translate these results into costs we augment our baseline coefficient target $\hat{\beta}$ to include the heterogeneity effects amongst non-participants presented in Table 7 and re-estimate the structural model allowing for heterogeneous costs of entry. Because continued participation effects are near zero for all groups, we do not allow for heterogeneity in per period costs of participation, but instead calibrate these to the median value of 200 SEK (30 USD).

Allowing for heterogeneous costs while maintaining computational feasibility requires a parametric assumption for the cost distribution. Denoting a vector of household characteristics x_i, we assume that χ_i is distributed truncated normal with mean $\mu_\chi x_i$, standard deviation σ_χ, and truncation at zero. Thus, the constant term in μ_χ represents the unconditional mean, while the term in μ_χ for each characteristic represents the shift in distribution for individuals with this characteristic. Results of this estimation are presented 10. For each characteristic, the coefficient of μ_χ an is presented on the same line as the characteristic it shifts. For instance, the coefficient of μ_χ associated with education shifts the mean for college educated individuals, and is presented on the same line as the effect for non-participants with a college education. The 4th column repeats the reduced form estimates, while the last column presents the fit of the estimated model.

Given the more rigorous structure placed on the cost distribution, it is not surprising that the model fit deteriorates. In general however the estimated model captures the main features of the data. The relative sign of the effect for each heterogeneity analyses is correct, although in many cases the model fails to generate as large of a difference. This is not surprising given that the entry cost distribution presented in Figure 4 is fairly polarized, and the assumed normal distribution can't replicate this. However, the model does do reasonably well in replicating the coefficients matched in the previous structural estimation. While the match is again not exact, the qualitative patterns are correct.

The implied costs generally reflect the findings of the reduced form analyses. The distribution of costs are estimated to have with constant mean of 1,232,000 SEK (180,000 USD) and with a standard deviation of 1,296,000 SEK (190,000 USD). College educated individuals are estimated to have a lower cost of entry on average of almost 500,000 SEK, reflecting that their participation probability is most affected by wealth. Although we don't explicitly model differences in beliefs, effects of recent equity returns are captured thorugh our estimate that winners following a year of positive stock returns have on average a lower cost of entry of over 200,000 SEK (30,000 USD). Similarly, although we don't model differences in labor income risk, self-employed individuals are estimated to have higher entry costs of over 400,000 SEK (60,000 USD). This estimate likely reflects their higher background risk. Demographically, older and female individuals are estimated to have higher costs, as are home owners. These patterns directly track the reduced form estimates. Interesting,ly, individuals with no debt are estimated to have larger costs of entry, although reduced form estimates suggest that such individuals are more likely to enter. This likely reflects differences in the distribution of state variables for this group, and highlights why one should exhibit caution in interpreting the reduced form heterogeneity analyses as evidence of higher costs. The estimated CDFs of entry costs are presented for each source of heterogeneity are presented in Figure 5. These again highlight the extremely high costs for all subsamples, as no group of non-participants is estimated to have a median entry cost of less than 1,000,000 SEK (150,000 USD).

Allowing for heterogeneous costs allows us to fit richer patterns of entry while still maintaining reasonable model fit. Although cost patterns vary intuitively along various characteristics, in all groups we find that extremely high costs of entry are necessary to match empirical effects. This confirms the main finding of our baseline structural estimation. In the next section we explore several possible reasons as to why estimated entry costs are so high, and how the disincentive to participate could modeled in a more believable manner.

5 Alternative Disincentives

Regardless of the exact model, the empirical estimates necessitate an extremely strong disincentive for pre-lottery non-participants to enter equity markets. As shown in the previous section, when the disincentive is restricted to financial costs the estimated costs of entry are enormous. Given that it is difficult to align an entry cost of over 2,700,000 SEK (400,00 USD) with any of the real financial costs necessitated by equity market entry, these costs seem absurd. Thus, it is natural to question what factors are omitted from the model or how these costs might be reinterpreted to provide more credible interpretations. In this section we do this in two ways. First, we conduct richer analysis of home ownership and how it is affected by wealth to see if equity market non-participation is explained by exposure to risky assets through real estate markets. We find no evidence to support this theory. Next, we examine whether non-participation can instead be explained by pessimistic beliefs of equity market returns. Our estimates suggest this to be a plausible channel.

5.1 Housing

Several prior studies, including Grossman and Laroque (1990) , Flavin and Yamashita (2002), Cocco (2005), and **?** have documented that equity market participation and real estate market particiaption should be considered jointly. Since returns on equities and real estate are correlated, owning both is less attractive. Furthermore, because home ownership provides substantial utility benefits but entails a large, lumpy expense, it is possible that an individual that purchases a home optimally chooses to hold no equity. However, if individuals choosing to purchase real estate instead of participating in equity markets explains the low estimated effect of wealth on equity market participation, than one would expect to see a significant increase in real estate holdings caused by receipt of wealth.

We first explore this in Figure 6. The left panel presents the unconditional effects of wealth on real estate wealth, as estimated from equation 1. This figure shows that exogneous wealth assignment does cause increases real estate wealth, suggesting that individuals do allocate some proportion of winnings to real estate. Although the effect is initally small, real estate wealth holdings rise to roughly 10% the amount of the lottery payout at longer horizons. Thus, households allocate a non-trivial proportion of wealth to real estate. To better understand this, the right panel presents the same estimated effect conditional on prior home ownership. Here, we see that households that owned real estate pre-lottery, there is no effect on real estate holdings. However, for individuals that did not own real estate pre-lottery, real estate holdings increase by roughly 20% of the amount won after three years. These patterns suggest that the unconditional effect of wealth on real estate is driven primarily through home purchase by pre-lottery non homeowners.

To understand what proportion and what type of winners purchase homes we define a partition of real estate holders and equity market participants and estimate a linear probability model to identify the effect that receiving 1M SEK has on the transition probability from one state to another. The results are presented in Table **??**. In the top panel we present the effect the year of the lottery. In the first column, we observer the effects of 1M SEK on the probability an transitions into owning neither real estate nor equity from various states. Not surprisingly, we estimate that for individuals that owned neither pre-lottery there is a decrease in the probability that they continue to own neither of .143. These individuals are thus choosing to purchase either equity or real estate. Other groups exhibit no effect of wealth on owning neither type of asset. In the second column we examine the effects on owning real estate. We observe that individuals that owned neither before are caused to have a higher probability of only owning housing of .040. We also observe that assignment of 1M SEK causes a decrease in probability of .097 pre-lottery households that only owned real estate continue in this category. Since such individuals were shown in column 1 to not estimated to have a higher probability of owning neither, these individuals must choose to participate in equity markets in some way.

In column 3 of Table **??** we present the estimated effect of wealth on households being in the only owning equity category. In the first row we estimate a .113 increase in probability of being in this category from households that did not own equity or real estate. In the third row of this column, the estimated effect suggests that households that were in this category pre-lottery are less likely to continue in this category. Since the previous two columns showed no effect on such individuals exiting equity markets, these households must be choosing to enter the real estate market. The third row of column 4 confirms this. In addition, in this column we observe that individuals that only owned housing before are .101 more likely to own both equity and housing after assignment of wealth, and individuals that owned neither have an improved probability of owning both of .029. The bottom panel of Table **??** exhibits similar patterns 3 years after the lottery. While we do observe a higher effect on the probability that households that owned neither own both types of assets at this horizons, other estimates are very close to estimates at $t = 0$.

Taken all together, these results suggest that while households that did not own real estate pre-lottery are estimated to have a higher probability of entering the real estate market, the effects are small. Households that owned neither neither real estate nor equity experience in increase of home ownership probability of .069 (.136), and households that owned equity only experience an increase of .057 (.056) in the probability of home ownership in some capacity at $t = 0$ ($t = 3$). Thus while housing is important, it does not explain the behavior of the majority of our sample. Columns 2 and 6 are key to understanding this, as they represent the state in which a household may choose to forgo equity market participation in favor of owning real estate. However, the effects in this column are estimated to be small. For households that enter this state non-participation in equity markets may be justified, but estimates suggest them to be a minority of the population.

Combining Figure 6 with Table **??**, we conclude that real estate effects are confined to a small number of individuals that use significant proportions of lottery winnings to purchase real estate. For such individuals real estate investment can explain non-participation. However, this group appears to be the minority of our sample, and do not appear to be driving our main results. If we were to incorporate real estate into the structural model, the lack of real estate and equity market entry would require strong disincentives to participate in both markets. Thus, modeling housing would not explain why households choose not to purchase equity, but rather create a further puzzle as to why households neither choose to purchase equity nor real estate. Therefore, real estate investment is not a plausible explanation as to why households in our sample are estimated to have only small increases in equity markets following receipt of large amounts of wealth.

5.2 Beliefs

We next explore the possibility that non-participation is caused by pessimistic beliefs of equity returns. Although historically the equity premium is large, it is possible that some households fail to accurately assess the benefits of equity ownership. Recent work by Van Rooij, Lusardi, and Alessie (2011) and Christelis, Jappelli, and Padula (2010) has emphasized that a lack of financial literacy contributes to non-participation in equity markets. Similarly, Guiso and Jappelli (2005) show that awareness of investment opportunities is an important determinant of stock market participation. Grinblatt, Keloharju, and Linnainmaa (2011b) show that that IQ is a significant predictor of stock market participation and portfolio efficiency. While these studies do not attempt to model these costs, they are generally interpreted as evidence that a lack of information, understanding, or attention contribute to the decision to not participate in equity markets. Furthermore, the stronger effect of wealth on more educate winners and winners following a year of positive equity returns provide hints that information, understanding, and beliefs contribute to the equity market participation decision.

Pessimistic beliefs lower the expected return to participation in equity markets, thus making participation less attractive. If a borrowing constraint is present, households wih low beliefs may optimally choose to forgo participation. To understand how this channel might explain the lack of participation for pre-lottery non-participants, in this section we re-estimate an alternate version of the structural model introduced in Section 4. In this version, we calibrate costs of participation and entry to zero. We then allow the expected equity premium, $\mathbb{E}[r_{s,t+1} - r_f]$ to differ from historical values and estimate the belief distribution that is consistent with the estimated effects of wealth on participation. Because our reduced form findings document no effect of wealth on continued equity market participation, we allow pre-lottery market participants to have beliefs consistent with historical returns. Thus, the estimated belief distribution reflects only the beliefs necessary to generate the participation responses of pre-lottery non-participants. The model is additionally augmented to include a no borrowing constraint to prevent individuals with pessimistic beliefs from short-selling equities and thus finding equity market participation to be attractive.

To implement belief heterogeneity, we suppose that household i solves the consumer problem under the belief that equity returns are distributed lognormally with mean $\bar{r}_i$ and variance σ_s, although in reality $\bar{r}^\star = .065$. Because heterogeneous $\bar{r}_i$ and variance σ_s are not separately identified, σ_s is assumed to be its historical average of .21. We choose not to permit learning or updating of beliefs, but beliefs regarding the equity return $\bar{r}_i$ can generally be thought of as the outcome of some rational inattention or informational choice problem. We assume that $\bar{r}_i \sim H_{\theta_{\bar{r}}}(\bar{r}_i)$. We parametrize this CDF similarly to the cost CDFs in Section 4.3. In particular, we construct a linearly spaced grid of 7 points denoted $x_{\bar{r}}^n$ between the risk free rate (r_f) and actual return ($\bar{r}^\star = .065$). Although in principle individuals could believe that the equity premium is negative, in this study's model with risk averse agents any belief below r_f would generate global non-participation, yielding beliefs below r_f unidentified. The CDF is assumed piecewise linear with CDF values defined

$$\theta_{\bar{r}} = \{\theta_{\bar{r}}^n | H_{\theta_{\bar{r}}}(x_{\bar{r}}^n) = \theta_{\bar{r}}^n\}. \tag{10}$$

Defining $\Theta = \theta_{\bar{r}}$, we then estimate equation 9 with the matched coefficients corresponding to those in the baseline estimation.

The results of this estimation procedure are presented in Table 12 and Figure 7. In Table 12 we present the model fit, and find that the model does reasonably well in matching the estimated effects of wealth on equity market participation. Although the model does overpredict the effect of moderate prizes on the participation and underpredicts the effect of large prizes on the participation of pre-lottery non-participants, it otherwise successfully matches the empirical estimates. In addition, it may be possible to improve the fit with a finer approximation of the CDF.

In Figure 7 the estimated distribution of pre-lottery non-participant beliefs is presented.The most striking feature of this distribution is that 53% of non-recipients are estimated to believe that the equity risk premium is negative. The estimated CDF further suggests that a small mass of individuals have beliefs that the equity premium is moderately positive, while approximately 32% of individuals have beliefs that are close to the estimated historical equity premium in Sweden. Thus, the estimated distribution assigns the majority of the distribution of equity return beliefs to be well below what historical returns would dictate.

Although the estimated belief distribution neccessary to replicate the estimated effects on participation is difficult to reconcile with past returns, previous work by Dominitz and Manski (2011) and Hurd, Van Rooij, and Winter (2011) has documented that subjective expectations of equity returns do not necessarily align with historical data. Dominitz and Manski (2011) use HRS to show that equity returns vary systematically with age, gender, and schooling, but remain remarkably stable over time. Hurd, Van Rooij, and Winter (2011) finds similar patterns in a survey of Dutch households, but notes that individuals end o be more pessimistic than would be estimated from recent stock market performance. Importantly, this study estimates separate distributions of subjective equity returns for equity market participants and non-participants. Because identification of the belief distribution in Figure 7 comes exclusively from non-participants, the estimated distribution of equity returns in Hurd, Van Rooij, and Winter (2011) for equity market non-participants provides a suitable benchmark to which we can compare our estimated distribution. In Figure 7 we therefore plot a second line that corresponds to the equity premium implied by non-participants used in Hurd, Van Rooij, and Winter (2011).[6]. In examining this second line, the first striking feature is that the calculated distribution of belliefs implies that 53% of Dutch non-equity market participants expect a zero or negative equity premium. This number is remarkably close to that implied by our estimated distribution (53%). In addition, the two distributions track very closely in the allowed range, rarely differing by more than .1. Although at the end our estimated CDF has a sharp uptick due t our assumption that expected equity premium could no assume the historical average, this figure demonstrates convincingly that the distribution of estimated beliefs necessary to replicate the participation responses we observe is consistent with subjective beliefs reported in surveys.

Although we can clearly not identify pessimistic beliefs of equity returns as the causal pathway to generate our estimated effects on participation, the similarity between our estimated distribution and self-reported subjective beliefs suggests that it this theory is plausible. More research on the formation of pessimistic beliefs might form and how this process interacts with wealth is needed to reject or validate this theory. However, given that such beliefs can replicate the estimated reduced form effects reasonably well, the theory is consistent with our main findings.

6 Conclusion

In this paper we have used random assignment of significant amounts of wealth to estimate the effect of wealth on stock market participation. Generally, we estimate a small, significant unconditional effects, a significant effect for individuals that did not own equity prior to assignment of wealth, and practically no effect for individuals that owned equity prior to assignment of wealth. Surprisingly, amongst individuals that did not own equity prior to wealth assignment, receiving 1M SEK has a fairly small effect on participation probability relative to prize size. In no specification, including amongst individuals that received more than 2M SEK (approximately 250k USD), do we observe the effect of wealth on participation probability rise above .325, and most specifications suggest an effect of less than .2. These effects suggest per annum costs of equity market participation above 18920 SEK (2823 USD) are necessary to explain non-participation for 67.5% of non-participants. This number not only is far different from what previous cross-sectional studies have found, but simply too large for recurring participation costs to be a valid explanation for observed non-participation. In addition, our estimates of no effect of wealth on participation for pre-assignment participants further suggests that participation costs are not suitable for modeling equity market participation.

In estimating our structural model, we allow for both costs of participation and costs of entry, and estimate the distribution of both using the method of indirect inference. We find costs of entry are much higher than have previously been found in other studies, with our estimated distribution implying median entry cost of more thatn 2,700,000 SEK (400,000 USD). Estimating a model that permits heterogeneous costs finds that low education, home ownership, being female, and self-employment are associated with higher costs of entry, while college education and receiving wealth following a year of positive equity return are associated with lower costs of entry. However, in both cases the magnitude of costs suggests that other mechanisms are likely driving non-participation.

Estimated costs of entry and participation likely reflect other channels that affect participation that aren't explicitly modeled. We explore two alternative channels that could account for the small estimated effects of wealth on equity market participation. First, we find that foregoing

[6] Note that Hurd, Van Rooij, and Winter (2011) fits the distributions of both mean and variance of subjective equity return expectations. We translate these numbers to expected equity returns, and assume a 2% risk free return to correspond to that of the current study.

equity market participation in favor of real estate purchases is unlikely to explain our estimated effects. Although exogenous assignment of lottery wealth does increase real estate holdings and real estate ownership probability, the estimated effects are too small for this to be a major contributing factor. Second, we find that pessimistic beliefs regarding expected stock returns plausibly explain the small estimated effects on participation. Modeling and estimating subjective beliefs can fit the empirical estimates decently, and, more importantly, results in estimated belief distributions that are consistent with beliefs measured in prior studies.

More broadly, our paper shows the dangers of using cross-sectional variation in characteristics to make inference regarding individual behavior. Identification from the cross-section is complicated by reverse-causality and unobserved heterogeneity, and we estimate extremely different implications from within individual variation of wealth than others have found when using the cross-section. Assumptions required to identify individual effects in cross-sectional studies are often quite strong, and the identification strategy should be carefully considered rather than taken at face value. Future work with this project will use our generated individual variation to examine not only the external margin effect of wealth on financial risk taking, but also the internal margin of adjustment.

In short, we utilize an extremely unique data set to estimate a behavior that has implications for several branches of economics. Our resulting estimates call into question a common belief that moderate costs of participation are capable of accounting for the lack of equity market participation, and suggest that future work should examine the interaction between wealth and expectations more carefully.

References

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

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

Barber, B. M., and T. Odean (2001): "Boys will be boys: Gender, overconfidence, and common stock investment," *Quarterly journal of Economics*, pp. 261–292.

Brunnermeier, M. K., and S. Nagel (2008): "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation," *The American Economic Review*, pp. 713–736.

Calvet, L., J. Y. Campbell, and P. Sodini (2009): "Fight or flight," *Portfolio rebalancing by individual investors, Forthcoming in The Quarterly Journal of Economics.*

Calvet, L. E., J. Y. Campbell, and P. Sodini (2007): "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115(5), 707–747.

Calvet, L. E., and P. Sodini (2014): "Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios," *The Journal of Finance*, 69(2), 867–906.

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

Cesarini, D., E. Lindqvist, R. Östling, and B. Wallace (2013): "Estimating the causal impact of wealth on health: Evidence from the Swedish lottery players," *Unpublished manuscript.*

Christelis, D., T. Jappelli, and M. Padula (2010): "Cognitive abilities and portfolio choice," *European Economic Review*, 54(1), 18–38.

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

Davis, S. J., F. Kubler, and P. Willen (2006): "Borrowing costs and the demand for equity over the life cycle," *The Review of Economics and Statistics*, 88(2), 348–362.

Dominitz, J., and C. F. Manski (2011): "Measuring and interpreting expectations of equity returns," *Journal of Applied Econometrics*, 26(3), 352–370.

Fagereng, A., C. Gottlieb, and L. Guiso (2013): "Asset market participation and portfolio choice over the life-cycle," *WP*, NA, NA.

Flavin, M., and T. Yamashita (2002): "Owner-occupied housing and the composition of the household portfolio," *American Economic Review*, pp. 345–362.

Friend, I., and M. E. Blume (1975): "The demand for risky assets," *The American Economic Review*, pp. 900–922.

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

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2011a): "IQ and Stock Market Participation," *Journal of Finance*, 66(6), 2121–2164.

——— (2011b): "IQ and stock market participation," *The Journal of Finance*, 66(6), 2121–2164.

Grossman, S. J., and G. Laroque (1990): "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica*, 58(1), 25–51.

Guiso, L., and T. Jappelli (2005): "Awareness and stock market participation," *Review of Finance*, 9(4), 537–567.

Haliassos, M., and C. C. Bertaut (1995): "Why do so few hold stocks?," *Economic Journal*, 105(432), 1110–1129.

Hankins, S., and M. Hoestra (2011): "Lucky in Life, Unlucky in Love? The Effect of Random Income Shocks on Marriage and Divorce," *Journal of Human Resources*, 46(2), 403–426.

Hankins, S., M. Hoestra, and P. M. Skiba (2011): "The Ticket to Easy Street? The Financial Consequences of Winning the Lottery," *Review of Economics and Statistics*, 93(3), 961–969.

Heaton, J., and D. Lucas (2000): "Portfolio choice and asset prices: The importance of entrepreneurial risk," *The journal of finance*, 55(3), 1163–1198.

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

Imbens, G. W., D. B. Rubin, and B. I. Sacerdote (2001): "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, 91(4), 778–794.

Kearney, M. S., P. Tufano, J. Guryan, and E. Hurst (2010): "Making savers winners: An overview of prize-linked savings products," Discussion paper, National Bureau of Economic Research.

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

King, M. A., and J. I. Leape (1984): “Wealth and portfolio composition: theory and evidence,” *Journal of Public Economics*.

Mankiw, N. G., and S. P. Zeldes (1991): “The consumption of stockholders and nonstockholders,” *Journal of Financial Economics*, 29(1), 97–112.

Merton, R. C. (1971): “Optimum consumption and portfolio rules in a continuous-time model,” *Journal of economic theory*, 3(4), 373–413.

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

Samuelson, P. A. (1969): “Lifetime portfolio selection by dynamic stochastic programming,” *The review of economics and statistics*, pp. 239–246.

Van Rooij, M., A. Lusardi, and R. Alessie (2011): “Financial literacy and stock market participation,” *Journal of Financial Economics*, 101(2), 449–472.

Van Rooij, M. C., A. Lusardi, and R. J. Alessie (2012): “Financial Literacy, Retirement Planning and Household Wealth*,” *The Economic Journal*, 122(560), 449–478.

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

Vissing-Jorgensen, A. (2002): “Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures,” Discussion paper, National Bureau of Economic Research.

Tables

Table 1: **Overview of Identification Strategy:** This table provides a brief summary of the idneification strategy used in each lottery. For a more detailed description please see Section 2.1.

	Period	Treatment	Cells
PLS	1990-1994	Total prize, fixed and odds	Prize draw × Balance
PLS	1995-2003	Total prize, fixed	Prize draw × No. wins
Kombi	1994-2008	Prize	Prize Draw × Balance
Tv-Triss	1997-2010	Prize	Year

Table 2: **Prize Distribution:** This table presents the distribution of prizes b size for the pooled sample, each of the lottery subsamples, and for the subsample of winners in year 1999 or later.

	Full Sample	**Kombi**	**TV-Triss**	**PLS**	**Post-1998**
$L_i \leq 100,000$	356,583	46,024	1,315	309,244	89,656
$100,000 < L_i \leq 500,000$	4157	0	2,532	1,625	1782
$500,000 < L_i \leq 1,000,000$	336	30	183	123	158
$L_i \geq 1000000$	991	432	220	339	529

Table 3: **Testing for Random Assignment:** This table presents results from estimating Equation 2 for the pooled sample, each of the lottery subsamples, and for the subsample of winners in year 1999 or later.

Variables	**Pooled**	**PLS**	**Kombi**	**TV-Triss**	**Post-1998**
Female	97.03	-177.274	1719.449	9858.218	3177.998
	(386.546)	(188.863)	(2475.357)	(24923.830)	(1298.907)
Age	-77.906	44.356	-4882.304	-183314.32	445.181
	(264.941)	(118.088)	(5105.629)	(17576.343)	(1019.945)
Age^2	2.843	-1.077	79.698	456.873	-4.595
	(4.923)	(2.242)	(86.630)	(347.121)	(18.413)
Age^3	-.022	.007	-.432	-3.178	.010
	(.028)	(.013)	(.475)	(2.148)	(.103)
Native	379.113	640.641	-1924.135	-10073.349	1911.564
	(1388.001)	(417.610)	(5080.025)	(56009.743)	(3747.967)
College Educated	20.717	-240.535	-2189.649	37132.564	-936.216
	(494.316)	(234.741)	(1423.011)	(33980.517)	(1273.356)
Married	-260.830	99.906	-1725.787	-14341.042	-1002.590
	(360.042)	(196.222)	(1629.589)	(34653.482)	(1045.088)
Income	-7.980e-4	-8.387e-4	.003	-.058	.006
	(1.778e-3)	(8.226e-4)	(.005)	(.096)	(.004)
$\hat{f}$	.70	.87	1.24	1.45	1.54
p-val	.67	.53	.28	.18	.15
N	362067	311331	46846	4250	92125
R^2	.0949	.038	.001	.007	.097

Table 4: **Comparison of samples:** This table compares our lottey sample to a population sample matched on age and gender. This is repeated for the subsample restricted to winners in 1999 or later.

Characteristic	Full Lottery (1)	Full Population (2)	Post-1998 Lottery (3)	Post-1998 Population (4)
Female	50.9	50.9	50.5	50.5
Age	58.8	58.7%	62.8	62.8
Household size	1.79	1.81	1.76	1.80
Income	203,008	185,793	200,838	190,798
Married	57.7%	56.9%	53.2%	53.0%
Retired	40.6%	40.0%	46.5%	43.7%
Student	2.4%	2.4%	1.9%	2.0%
Self Employed	4.6%	3.9%	4.0%	4.8%
Native	97.2%	93.4%	97.2%	93.0%
College	18.3%	16.6%	19.8%	20.4%
Total Wealth			1,058,013	946,965
Total Debt			242,121	264,933
Own Home			67.5%	60.8%
Participate (Stock or MF)			67.4%	62.4%
Participate (Stock only)			42.2%	37.8%

Table 5: **Probit Estimation of Participation:** This table is a replication of Table 6 in **?**. The specified probit regression is estimaed on both our lottery and matched population sample, the results from the replicated table are presented for comparison. Marginal effects are calculated as the effect of increasing each variable by one standard deviation or by changining an indicator variable to one when all other variables are set to their median level.

Variables		**Lottery**			**Matched Population**			**CCS 2007**	
	Estimate	**z**	**Change**	**Estimate**	**z**	**Change**	**Estimate**	**z**	**Change**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Disposable Income	8.7e-7	4.27e-7	6.02%	3.3e-5	10.39	2.16%	10.6	15.30	8.80%
Private Pension Premia/Income		—			—		3.053	14.00	6.10%
Financial Assets (log)	.264	71.16	21.88%	.181	50.92	26.19%	.491	115.00	22.9%
Total Real Estate (log)	.010	1.38	1.93%	.035	6.66	11.36%	.021	22.80	3.3%
Total Liabilities (log)	.019	2.80	3.49%	-.021	-3.82	-3.02%	.017	13.50	4.6%
Retired	.023	.72	0.74%	.065	2.59	1.23%	.137	6.04	-3.8%
Entrepreneur	.011	-.35	0.61%	.046	1.49	1.28%	-.030	-.94	-1.0%
Unemployed	.084	2.32	2.63%	.068	2.60	0.86%	-.065	-3.61	-2.3%
Student	.065	1.29	2.05%	.109	16.37	3.67%	.028	1.08	1.0%
Age	-.012	-11.15	-5.95%	-.009	-12.08	-3.15%	-.014	-26.30	-9.8%
Household Size	-.029	-2.74	-3.05%	.112	1.79	.34%	-.009	-1.62	-.4%
High School education	.121	5.22	3.76%	.113	6.38	2.18%	.195	13.50	7.1%
Higher Degree	.214	7.82	6.41%	.246	12.66	5.15%	.130	9.12	4.3%
Missing Education	.182	2.29	5.07%	.491	5.31	6.72%	-.066	-26.30	-2.3%
Immigrant	-.092	-2.61	-0.95%	-.268	-9.19	-4.07%	-.384	-26.30	-14.4%
Constant	-2.01	-27.15		-3.74	-39.61		-4.852	-98.80	

Table 6: **Baseline Regressions Regression:** This table provides results from the baseline regrssions reported in the baseline empirical analyses.

	Stock or Mutual Fund				Stock Only	
	Unconditional	**Event Study**	**Conditional**	**Threshold**	**Unconditional**	**Conditional**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
L_i	.038	.043			.020	
	(.010)	(.007)			(.012)	
$L_i \times Part_{-1}$			.003			.020
			(.004)			(.007)
$L_i \times NonPart_{-1}$			.113			.007
			(.023)			(.011)
$L_i \times Part_{-1} \times (100k < L_i < 1M)$				.003		
				(.014)		
$L_i \times Part_{-1} \times (1M < L_i < 2M)$				.005		
				(.012)		
$L_i \times Part_{-1} \times (2M < L_i)$				-.006		
				(.030)		
$L_i \times NonPart_{-1} \times (100k < L_i < 1M)$				.060		
				(.021)		
$L_i \times NonPart_{-1} \times (1M < L_i < 2M)$				.185		
				(.044)		
$L_i \times NonPart_{-1} \times (2M < L_i)$				.325		
				(.084)		
$Female$	-.038	-.040	-.004	-.003	-.067	-.007
	(.004)	(.003)	(.002)	(.002)	(.005)	(.002)
Age	-.019	-.006	-.001	-.001	.019	.004
	(.003)	(.002)	(.001)	(.001)	(.003)	(.001)
Age^2	5.48e-4	3.21e-4	2.33e-5	2.35e-5	-5.87e-5	-6.77e-5
	(5.65e-5)	(3.36e-5)	(2.24e-5)	(2.25e-5)	(5.66e-5)	(2.36e-5)
Age^3	-4.44e-6	-3.28e-6	2.15e-7	2.16e-7	-1.25e-6	3.04e-7
	(3.32e-7)	(2.05e-7)	(1.31e-7)	1.31e-7	(3.29e-7)	(1.34e-7)
$Native$	.091	.081	.005	.005	.035	.003
	(.012)	(.008)	(.005)	(.005)	(.013)	(.005)
$College_{-1}$	.116	.110	.008	.008	.155	-.010
	(.004)	(.003)	(.002)	(.002)	(.005)	(.002)
$Wealth_{-1}$			6.45e-9	6.45e-9		7.80e-9
			(3.75e-10)	(3.75e-10)		(4.82e-10)
N	79789	775043	73658	73658	79789	73658
R^2	.065	.071	.840		.068	.744

Table 7: **Heterogeneous Effects:** This table presents the coefficients of wealth interacted with non-paricipation in the year prior to the lottery for each listed caraceteristic. We omit the full set of coefficients which include a cubic in age, immigrant status, education the year prior to the lottery, and financial wealth the year prior to the lottery.

	Home Owner		**Debt**		**Self-Employed**		**Prior Equity Returns**	
	Yes	**No**	**$\leq$ 0**	**$>$ 0**	**Yes**	**No**	**$\leq$ 0**	**$>$ 0**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Effect	.104	.142	.211	.093	.031	.123	.055	.140
s.e.	(.027)	(.049)	(.036)	(.029)	(.024)	(.025)	(.039)	(.028)
p-val	.007	.000	.000	.000	.192	.000	.162	.000
N	9745	10543	9719	10569	509	19779	12308	7980
p-val equal effects	.499		.008		.009		.077	

	Gender		**Age**		**College Educated**	
	Male	**Female**	**$\leq$ 45**	**$>$ 45**	**Yes**	**No**
	(9)	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Effect	.147	.112	.156	.107	.230	.098
s.e.	(.038)	(.030)	(.045)	(.029)	(.052)	(.026)
p-val	.000	.000	.001	.000	.000	.000
N	9251	11037	2237	18051	2400	17888
p-val equal effects	.525		.367		.023	

Table 8: **Structural estimation results:** The resulting coefficiens from repeating the reduced form estimation on the data set generated by the structural model. Column 2 does this when the model is calibrated to the cross-section implied participation costs, while column 3 presents the cofficients resulting from the baseline structural estimation.

Effect	**Reduced Form Estimate (1)**	**Calibration (2)**	**Baseline Estimation (3)**
Unconditional	.038	.090	.044
Pre-existing Wealth	6.45e-9	.438e-9	2.07e-8
P	.003	.057	4.41e-4
NP	.112	.261	.127
100k-1M (P)	.003	.087	-.001
1M-2M (P)	.005	.175	.014
2M+ (P)	-.006	.155	-.016
100k-1M (NP)	.060	.107	.056
1M-2M (NP)	.185	.437	.181
2M+ (NP)	.325	.438	.338

Table 9: **Costs in Literature:** This table provides a brief overview of prior studies in the literature, including their estimated costs.

	Participation Cost	Entry Cost	Risk Aversion Aversion	Housing Modeled	Source of Identification	Other Comments
Vissing-Jorgensen (2002)	$250	—	—	—	Cross-Section	4% certainty equiv. risk premium
Andersen and Nielsen (2011)	1550 USD	—	—	—	Quasi-Experimental (Inheritance)	
Gomes and Michaelides (2005)	—	2.5% Income (750-1300 USD)	5	Yes	Cross-Section	Epstein-Zin; No Borrowing No Return to Housing
Cocco (2005)	1000 USD	—	5	Yes	Cross-Section	
Alan (2006)	—	2.15% (650-1100 USD)	2	No	Cross-Section	No
Khorunzhina (2013)	4-6% 1200-3050 USD	—	2-3	No	Panel	Solved Biannually $\beta = .9$
Fagereng, Gottlieb, and Guiso (2013)	120 USD	—	12	No	Panel	Finnish Administrative Records, $\beta = .89$

Table 10: **Heterogeneous Structural Estimation:** This table provides the estimates and resulting coefficients that are estimated from a structural model that permits heterogeneity in entry cost distributions along the specified dimensions. Note that this estimation finds $\sigma_{chi} = 1295749$.

Characteristic	μ_χ **(in SEK)**		**Reduced Form Estimate**	**Structural Estimation**
Unconditional	(1,231,999.5)		.038	.045
Pre-existing Wealth			6.45e-9	1.45e-8
P			.003	.005
NP			.112	.124
Prize Size		**100k-1M (P)**	.003	.013
		1M-2M (P)	.005	.005
		2M+ (P)	-.006	.026
		100k-1M (NP)	.060	.036
		1M-2M (NP)	.185	.222
		2M+ (NP)	.325	.267
Home Owner	(103,521)	**Yes**	.104	.115
		No	.142	.185
Debt	(10,245)	≤ 0	.211	.189
		> 0	.093	.126
Self-Employed	(415,493)	**Yes**	.031	.040
		No	.123	.145
Prior Equity Returns	(-205,836)	> 0	.140	.140
		≤ 0	.055	.112
Gender	(80,562)	**Female**	.112	.082
		Male	.147	.204
Age	(141,886)	> 45	.107	.106
		≤ 45	.156	.132
College Educated	(-491385)	**Yes**	.230	.245
		No	.098	.126

Table 11: **Real Estate:** This tabl presents the effects of wealth on real estate ownership and equity market participation status. Each number in a given row represents the effect that wealth has on he probability that a housheold enters the state specified in a column. This is repeated at horizons $t = 0$ and $t = 3$ respectively.

	At Time t=0 Own			
At Time t=-1 Own	**Neither Real Estate nor Equity (1)**	**Real Estate Only (2)**	**Equity only (3)**	**Real Estate and Equity (4)**
Neither Real Estate nor Equity ($N = 9,745$)	-.143 (.052)	.040 (.032)	.113 (.044)	.029 (.020)
Real Estate Only ($N = 10,543$)	-.011 (.011)	-.097 (.027)	.001 (.004)	.101 (.027)
Equity Only ($N = 14,051$)	-.008 (.013)	.006 (.009)	-.050 (.036)	.051 (.032)
Real Estate and Equity ($N = 39,319$)	.002 (.002)	-.001 (.005)	.008 (.007)	-.011 (.009)

	At Time t=3 Own			
At Time t=-1 Own	**Neither Real Estate nor Equity (5)**	**Real Estate Only (6)**	**Equity only (7)**	**Real Estate and Equity (8)**
Neither Real Estate nor Equity ($N = 7,230$)	-.125 (.070)	.055 (.047)	.020 (.03)	.081 (.049)
Real Estate Only ($N = 7,872$)	-.015 (.011)	-.045 (.045)	-.003 (.007)	.108 (.036)
Equity Only ($N = 11,445$)	.022 (.035)	-.019 (.008)	-.083 (.042)	.075 (.04)
Real Estate and Equity ($N = 31,628$)	-.002 (.002)	-.024 (.006)	.010 (.011)	.006 (.012)

Table 12: **Structural Estimation, Beliefs:** This table presents the model fit of the structural estimation of the subjective belief distribution.

Effect	Reduced Form Estimate	Belief Estimation
Unconditional	.038	.035
Pre-existing Wealth	6.45e-9	1.69e-8
P	.003	-2.01e-4
NP	.112	.109
100k-1M (P)	.003	2.38e-5
1M-2M (P)	.005	.013
2M+ (P)	-.006	-.018
100k-1M (NP)	.060	.100
1M-2M (NP)	.185	.198
2M+ (NP)	.325	.287

Figures

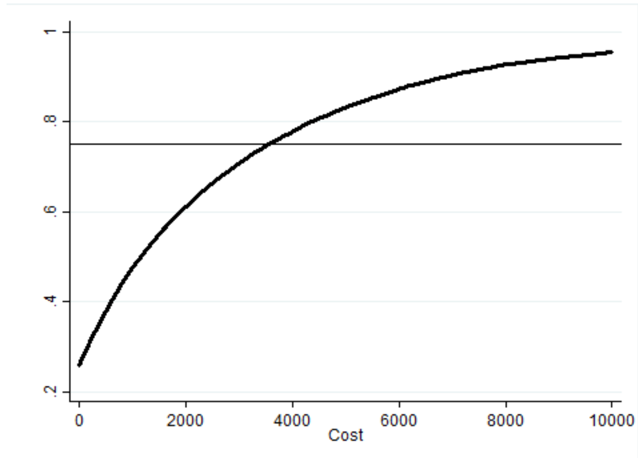

Figure 1: **Replication of estimation Vissing-Jorgensen (2002), Estimation A:** This figure presents the distribution of participation costs amongst non-participants suggested form our 1999 cross-section of lottery winners.

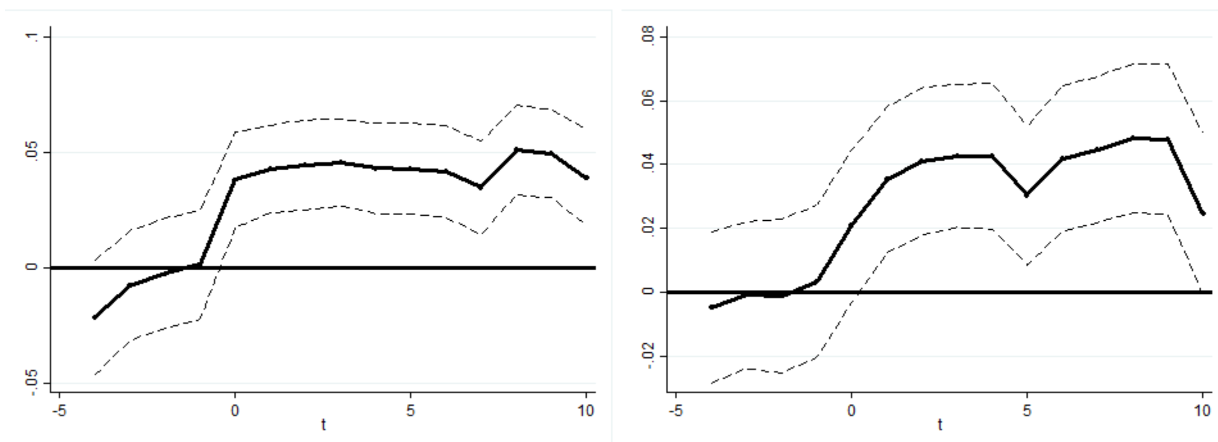

Figure 2: **Unconditional Effect:** Panel A presents the unconditional effect of wealth on equity market participation probability when participation is defined as owning stocks or mutual funds. Panel B presents the unconditional effect of wealth on equity market participation probability when participation is defined only as owning stocks.

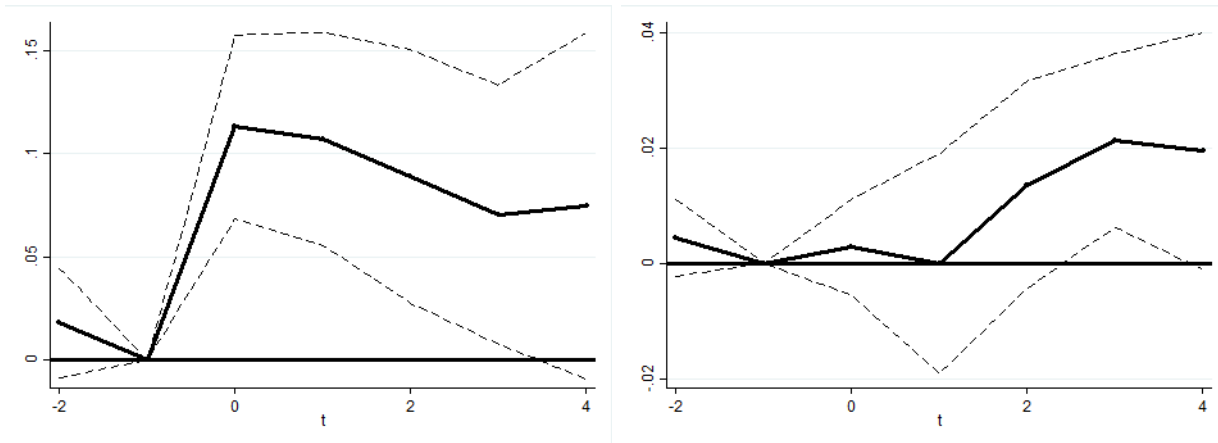

Figure 3: **Conditional Effects:** Panel A presents the effect of wealth on participation for pre-lottery non-participants. Panel B presents the corresponding effect of wealth on participation for pre-lottery participants.

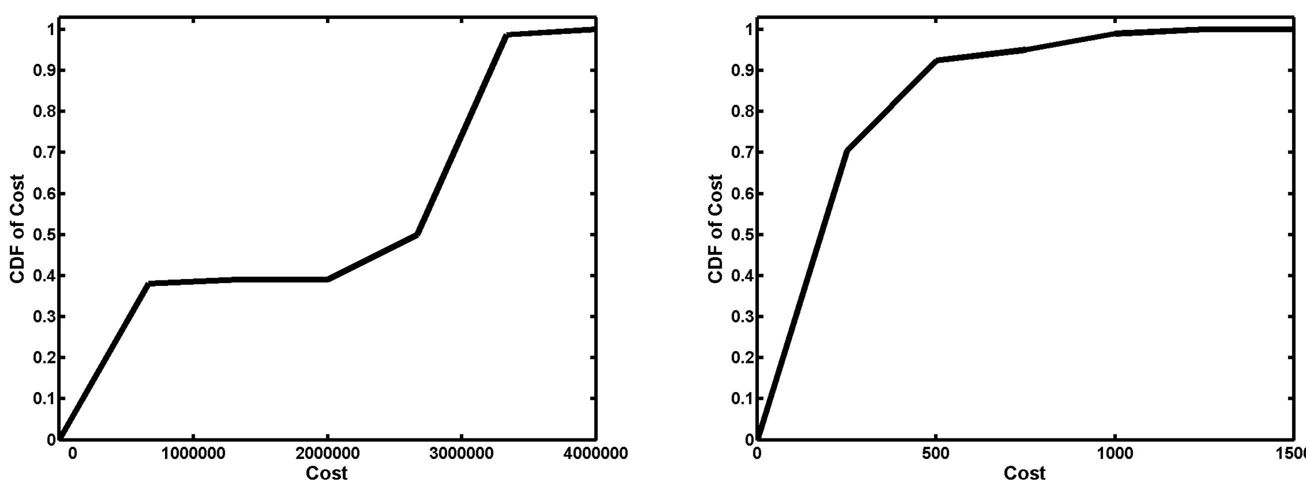

Figure 4: **CDF Estimates, Baseline Structural Estimation:** Panel A presents the estimated entry cost distribution, Panel B presents the estimated participation cost distibution.

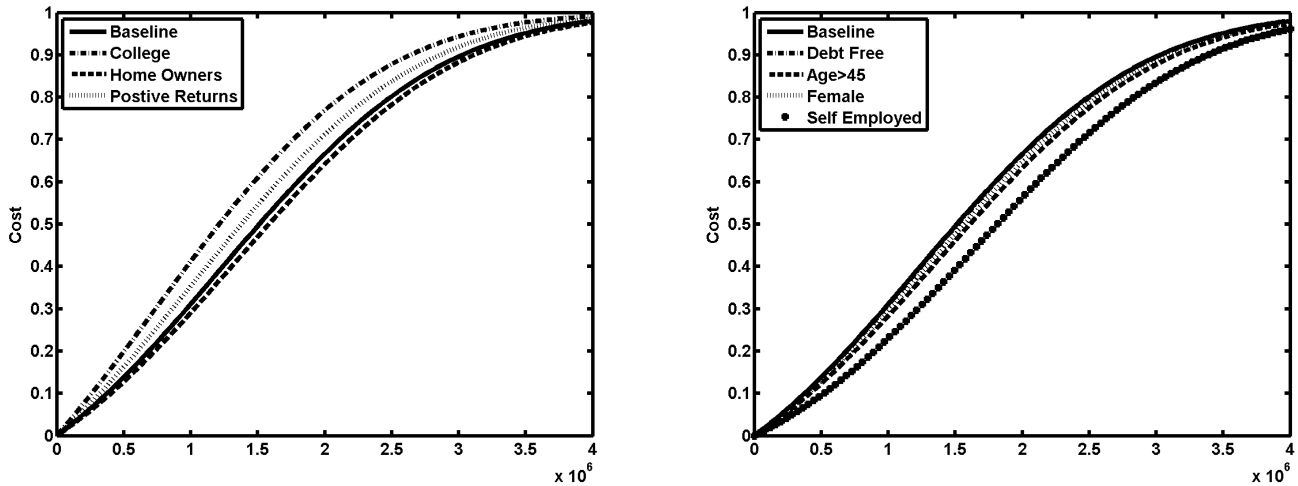

Figure 5: **CDF Estimates, Heterogeneous Structural Estimation:** Panel A presents the estimated entry cost distribution and the effect of college education, home ownership, and winning following a year of positive returns. Panel B A presents the estimated entry cost distribution and the effect of being debt free, over age 45, female, and self employed.

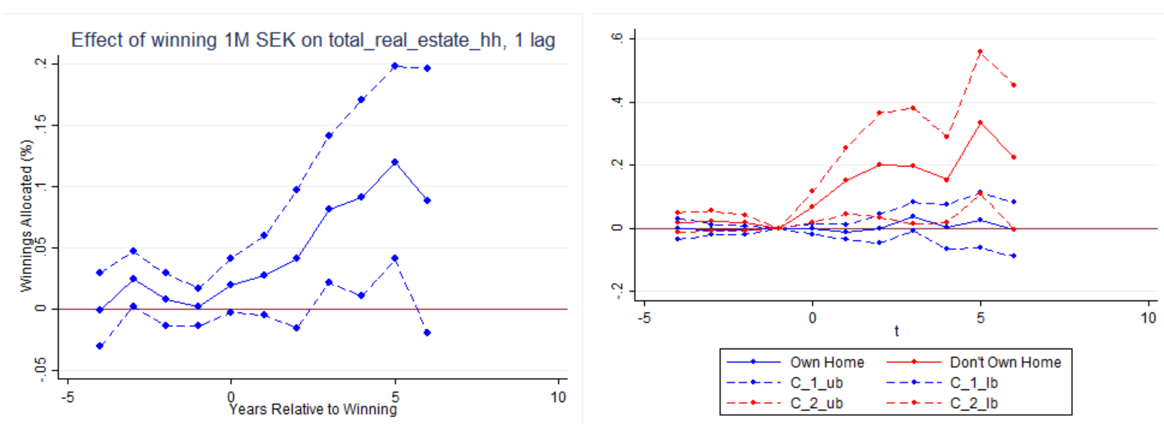

Figure 6: **Real Estate Effects:** Panel A presents the estimated unconditional effect of receiving 1M SEK on real estate holdins. Panel B presents the effect contingent on pre-lottery real estate holding status.

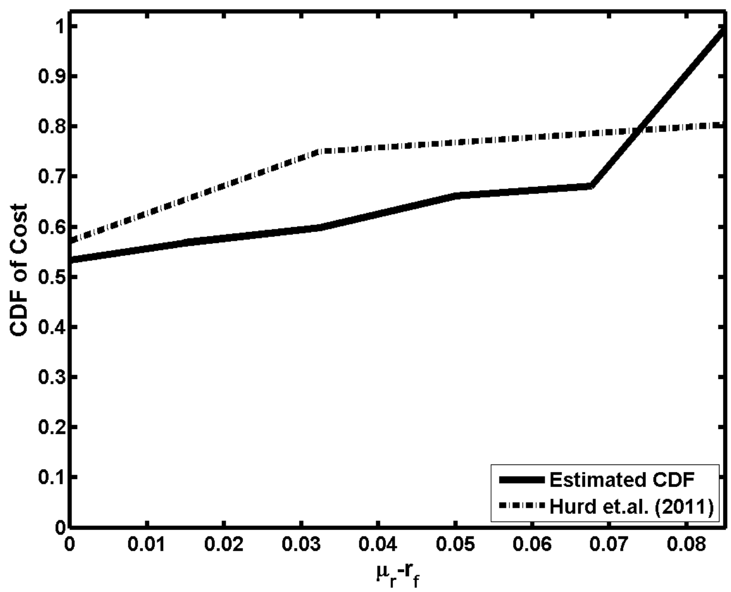

Figure 7: **CDF Estimates, Structural Estimation of Beliefs:** This figure presents the estimated distribution of beliefs, and compares this distribution to the measurement of beliefs done in Hurd, Van Rooij, and Winter (2011)

A Validation and Comparison

A second estimation of fixed costs of participation in Vissing-Jorgensen (2002) allows for population heterogeneity under the assumption that costs are uncorrelated with wealth. Assuming participation is optimal iff

$$W_{i,t} \times \alpha_i \times (r_{ce,t} - r_f) \geq F_i^P \tag{11}$$

a non-parametric estimate of the distribution of participation costs (denoted G_{F^P}) is the percentage of households with a given wealth level that participate in the stock markets. For example, if a given wealth level $\bar{W}$ implies a participation benefit $Benefit_t(\bar{W})$, then

$$G_{F^P}(Benefit_t(\bar{W})) = \frac{\#\ participants\ with\ W = \bar{W}}{\#\ of\ individuals\ with\ W = \bar{W}}. \tag{12}$$

To implement this, Vissing-Jorgensen (2002) divides the sample into ten deciles and plots the implied CDF against the median of each wealth decile (see figure 7). We replicate this exercise in figure 8. Here, we see that the median participation cost is XXX and 75% of non-participants are estimated to have participation costs lower than XXX SEK (XXX USD). Again, this estimation applied to our Swedish sample supports the finding of Vissing-Jorgensen (2002) and others that moderate participation costs are capable of accounting for the majority of non-participation.

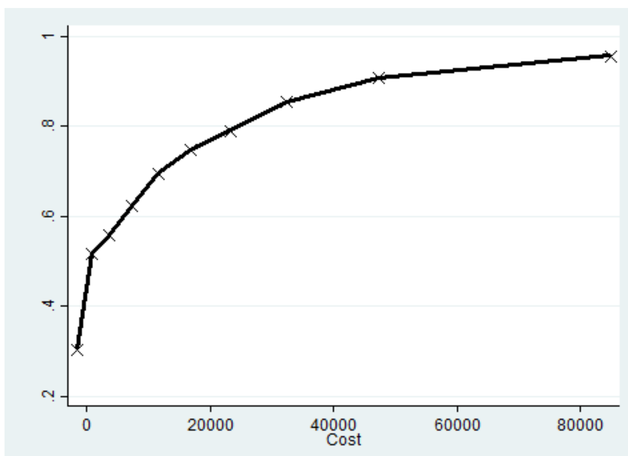

Figure 8: **Replication of Vissing-Jorgensen (2002), estimation B:** This figure presents the distribution of participation costs from the 1999 cross-section of lottery winners.

B Data

See Cesarini, Lindqvist, Östling, and Wallace (2013) for more detail. To be added here before posting online.

C Robustness exercises

C.1 Household Definition

Table 13: **Household Definition:** This table presents the estimated effects when the household definition is restricted to include only the winner.

	At Time t=0 Own		
	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	.043	.004	.101
SE	(.008)	(.004)	(.021)

C.2 Direct Participation only Results

Table 14: **Direct Participation:** This table presents the estimated effects when the participation is limited to owning equities directly.

	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	.020	.020	.007
SE	(.012)	(.007)	(.011)

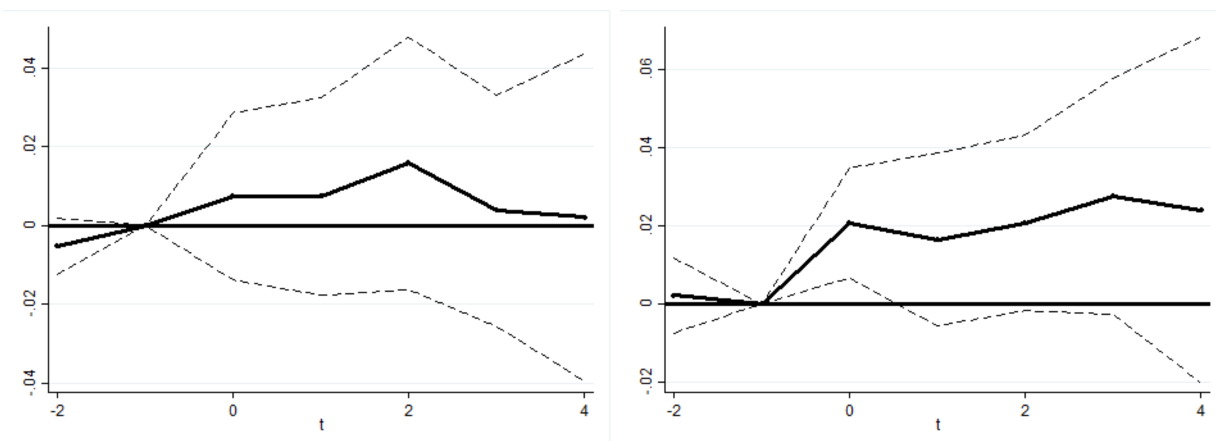

Figure 9: **Conditional Effects:** Panel A presents the effect of wealth on participation for pre-lottery non-participants. Panel B presents the corresponding effect of wealth on participation for pre-lottery participants. Here participation statusis defined as directly owning stock.

C.3 Separate Lotteries

C.4 Alternative Estimators

Table 15: **Kombi:** This table presents the estimated effects for the Kombi lottery subsample.

	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	.049	.007	.117
SE	(.014)	(.007)	(.032)

Table 16: **Triss:** This table presents the estimated effects for the Triss lottery subsample.

	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	.027	.001	.069
SE	(.011)	(.005)	(.025)

Table 17: **PLS:** This table presents the estimated effects for the PLS lottery subsample.

	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	-.012	.005	-2.287
SE	(.023)	(.010)	(2.059)

Table 18: Probit

	At Time t=0 Own		
	Unonditional	**Conditional on P**	**Conditional on NP**
Effect	.114	.062	.426
SE	(.035)	(.084)	(.092)
Marginal Effect			

D Technicals

D.1 Model Solution

Model solution algorithm summary and model code are both available on request.

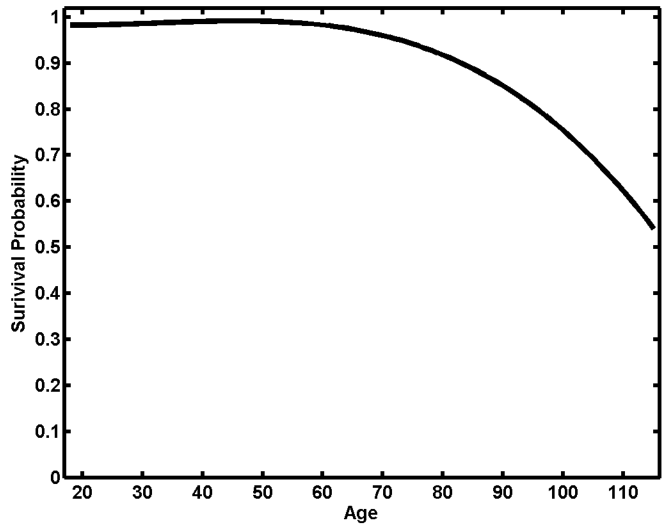

Figure 10: **Survival Probabilities:** This figure presents the estimated one year survival probability for each age.

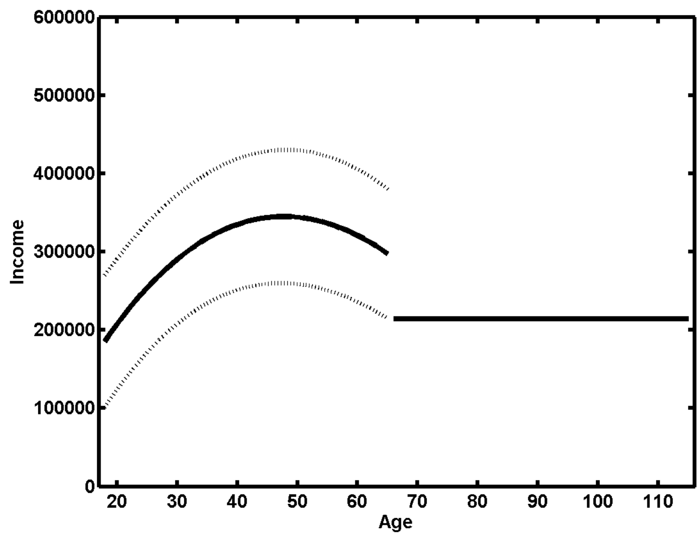

Figure 11: **Income Profiles:** This figure presents the estimated income for each age. For ages below retirement, it alsoshows the high and low transitor income components.

D.2 Estimation Technicals