The Effect of Education on Equity Holdings

Dmytro Hryshko[1] **María José Luengo-Prado**[2]
Bent E. Sørensen[3]

February, 2012

Abstract

We study the effect of education on equity ownership in the form of stocks or mutual funds (outside of retirement accounts). We find a causal effect of education on stockholding using the number of colleges in the county where the respondent grew up as an instrument and data from the Panel Study of Income Dynamics. The effect is particularly strong for whites from non-privileged backgrounds. We explore the channels through which education affects equity holdings using the Wisconsin Longitudinal Survey and find that, controlling for family fixed effects, increased cognition and features associated with having a white collar job appear to be the main channels.

For useful comments, the authors thank Eric Zitzewitz and two very thorough referees as well as participants at the 2008 PSID conference on "Pensions, Private Accounts, and Retirement Savings over the Life Course." [1]University of Alberta, 8-14 HM Tory Building, Edmonton, AB, T6G 2H4, Canada. Tel.: (780) 492-2544; fax: (780) 492-3300; e-mail: dhryshko@ualberta.ca. [2]Northeastern University, 301 Lake Hall, Boston, MA 02115-5000, USA. Tel.: (617) 373-4520; e-mail: m.luengo@neu.edu. [3]University of Houston and CEPR, 204 McElhinney Hall, Houston, TX 77204, USA. Tel.: (713) 743-3841; e-mail: bent.sorensen@mail.uh.edu

1 Introduction

The number of households participating in the stock market in the United States and other advanced economies does not exceed 50% in spite of stocks delivering higher average returns than other assets. Haliassos and Bertaut (1995) suggest that a "fixed cost" of participation is the main reason why participation is not higher. "Fixed costs" may be pecuniary, such as a fixed fee for opening a brokerage account, or non-pecuniary, such as ignorance of stocks or behavioral biases which may prevent investors from properly evaluating expected gains from stockholding—see Campbell (2006) for further discussion. It appears that non-pecuniary costs play a significant role because monetary costs of investing in the stock market have declined steeply over time while the share of (non-retirement) stockholders has remained roughly constant (see Table 1). Education, which is likely to lower non-pecuniary costs of stockholding, may therefore be vital for spreading stock ownership.

We demonstrate that higher education is associated with higher propensity to own stocks using data from the Panel Study of Income Dynamics (PSID). A high correlation between education and equity ownership need not reflect a causal relation because unobserved variables such as household attitudes, preferences, and abilities are likely to affect both the propensity to own equity and the choice of education. We uncover the causal effect of education on household participation in the stock market using instrumental variable (IV) techniques. As an instrument for education, we use the number of 4-year colleges when the respondent was 17 in the county where he or she grew up—an instrument used previously by Currie and Moretti (2003).[1] The identifying assumption is that the number of 4-year colleges in the county is orthogonal to unobserved household traits and other factors that may determine future propensities to hold stocks. An example of such a factor could be peer effects if college-rich counties are also counties with many stockholders.[2]

College availability will potentially lower the cost of college education. This can be due to a reduction in the distance between a college of choice and the family's residence which may allow students to live at home and save on the cost of shelter or it may make it easier for students to hold a part-time job near the college. It may also be that increased competition between colleges makes them recruit more aggressively, for example by lowering fees or

[1] We are grateful to Janet Currie for sharing the data with us.

[2] Peer effects, which we discuss in more detail in the following, could work through parents interacting with stockholding neighbors followed by a likely spillover from parental stockholding to children's stockholding; or it could be that household heads growing up in counties with high density of colleges and stockholders tend to stay in such counties.

offering more financial aid. For students who need to finance college mainly by borrowing, or students who face relatively high costs of borrowing during college years, a lower cost of college will make it more attractive to obtain a college degree—for a formal model, see Cameron and Taber (2004). In the PSID, the effect of college availability is stronger for heads who grew up in less advantaged families, which is consistent with the cost of college education being an important determinant of college completion. We find a strong and significant effect of education on the incidence of household stockholding with a stronger effect for white individuals who report growing up in non-privileged households. "Stockholding," when we analyze PSID data, refers to holdings of stocks and mutual funds outside of retirement accounts.

Having established a causal effect of education on stockholding using the PSID, we use the Wisconsin Longitudinal Survey (WLS) to examine potential mechanisms. We confirm, using ordinary least squares (OLS), the strong correlation between a college degree and non-retirement holding of equity (which includes bonds in the analysis of the WLS data). We then, one-by-one, include variables that are likely to be impacted by going to college, such as holding a white collar job. If the inclusion of a set of variables renders the college dummy insignificant, these variables, or unobserved variables highly correlated with them, must be the channels through which college affects equity holding. The WLS is well suited for this task because of the rich information collected—for example, siblings are interviewed—which allows us to control for family background using family-fixed effects. The WLS dataset is also quite unique, having a measure of IQ at the time the respondent attended high school and a measure of cognition obtained after college.

Increased cognition and holding a white collar job appear to be the most important channels through which college affects the propensity to own stock. Our two-pronged approach is similar to the strategy followed by Cole and Shastry (2009). They examine financial market participation using a large U.S. Census dataset, where they identify a household as participating in financial markets if the household reports having dividend, interest, or rental income. They use state-variation in compulsory schooling as an instrument and explore, with no instruments, the channels through which education may affect financial market participation using detailed survey data from the National Longitudinal Survey of Youth (NLSY). The IV results of Cole and Shastry (2009) are driven by the change in financial behavior of those who are stimulated by compulsory schooling laws to finish high school and who otherwise would not have done so, and the results therefore capture a so-called Local Average Treatment Effect (LATE) on these individuals. We similarly capture a LATE, namely the change in financial behavior of those who are stimulated

by college availability to obtain a college degree, but our results differ in that we are able to examine stock holdings while Cole and Shastry (2009) are limited to broader asset holdings which include savings accounts and bonds in addition to equity.

Our paper seems to be the first paper that examines the relation between stockholding and education in the United States using instrumental variables. A related paper using instrumental variables estimation is Christiansen, Joensen, and Rangvid (2007) who demonstrate that a university economics education has a causal positive effect on the propensity to own stock in Denmark.

The rest of the paper is organized as follows. Section 2 is a review of the relevant literature. In Section 3, we describe the PSID data and present Probit and IV-Probit results. Section 4 describes the WLS and presents Probit and OLS regressions with and without family fixed effects, while Section 5 concludes.

2 Literature Review

Many households in the United States choose not to participate in the stock market. Understanding this non-participation is important for several reasons. First, there is substantial heterogeneity in household wealth holdings after controlling for household demographic characteristics and income (e.g., Campbell 2006). Low participation rates in the stock market may help explain part of this heterogeneity as households who have stocks in their portfolios may end up with higher wealth (compared to otherwise similar households) because stocks offer, on average, higher returns than other assets. Low participation rates can also result in low wealth holdings for the average household at the time of retirement. Furthermore, low participation in the stock market can contribute towards an explanation of the equity premium puzzle—the inability of the consumption-based asset pricing model to reconcile the low correlation between aggregate consumption growth and stock excess returns with plausible levels of risk aversion. Indeed, as suggested by several studies (e.g., Attanasio, Banks, and Tanner 2002, Vissing-Jørgensen 2002a), what may matter for the determination of stock prices and excess returns is the consumption risk of those who participate in the stock market.

Why does education affect the probability of participating in the market for risky assets? Haliassos and Bertaut (1995) highlight inertial factors while Alan (2006) and Gomes and Michaelides (2005) show that small fixed costs of participating in the stock market—such as the time spent gathering

and processing relevant information—can help account for non-participation in dynamic life-cycle models.[3] Perraudin and Sørensen (2000) find, using the U.S. Survey of Consumer Finances, that observed zero holdings of stocks and bonds by many U.S. households can be attributed to fixed and proportional utility costs of holding non-zero amounts of these assets—costs that are higher for the least educated. Andersen and Nielsen (2010) show that a significant fraction of individuals receiving large unexpected inheritances eschew stock ownership—even to the extent of selling off inherited stock—a pattern which strongly indicates fixed costs of a non-monetary form. Education may lower these fixed costs to the extent that such costs reflect a lack of understanding of or familiarity with stocks.

Earnings risk ("background risk") in the form of high variance of permanent income may make it less attractive to hold risky assets such as stocks. Guiso, Jappelli, and Terlizzese (2005) provide evidence that such background risk affects portfolio choice—and because background risk may correlate with education we verify that our results are robust to the inclusion of measures of income risk.

More schooling can raise the likelihood of households' participation in the stock market by enhancing cognition or mathematical ability of students. Likewise, college education may increase financial literacy and knowledge about financial products, such as stocks and mutual funds. van Rooij, Lusardi, and Alessie (2011) directly measure financial literacy and find that it has strong predictive power for stock market participation. Financial literacy may itself be a function of individuals' IQ and Grinblatt, Keloharju, and Linnainmaai (2011), using a unique dataset from Finland where males' IQ is measured upon induction in mandatory military service, find a strong correlation of IQ with stockholding. Hansen, Heckman, and Mullen (2006) show that individuals with more schooling achieve higher scores on standardized tests of individual mathematical reasoning.[4] Christelis, Jappelli, and Padula (2010) find that general cognitive ability predicts stockholding using data from the Survey of Health, Ageing and Retirement in Europe, as do Cole and Shastry (2009) using NLSY data. Cognitive ability may be a function of education or it may be a function of innate IQ—probably both. van Rooij, Lusardi, and Alessie (2011) find that in regressions which include measures of education or measures of

[3] In the life-cycle models of Alan (2006) and Gomes and Michaelides (2005) participation costs are assumed to be proportional to permanent income, and they are best interpreted as the time (opportunity) cost of learning about the stock market instead of working.

[4] Calvet, Campbell, and Sodini (2007), using a comprehensive data set from the Swedish tax authorities, argue that it may be rational for some low-ability individuals to stay out of the stock market anticipating likely investment mistakes.

ability together with financial literacy, the latter is significant while the others are not, which indicates that the effect of IQ and education to a large extent works through financial literacy.

Attending college may impact the propensity to hold stock through peer effects as indicated by the results of Brown, Ivkovic, Smith, and Weisbenner (2008). This may occur through mimicking or because peers, such as neighbors and coworkers, talk about their earnings in the stock market. van Rooij, Lusardi, and Alessie (2011) show that individuals with highly educated acquaintances are more likely to hold stock while individuals with less educated acquaintances are not. One's acquaintances are not exogenously assigned and educated individuals are not necessarily financially literate, but this nonetheless indicates that knowledge imparted from informed peers may be important. Hong, Kubik, and Stein (2004) find that "more sociable" households who know their neighbors and frequently attend church are more likely to participate in the stock market. Sociability can also be affected by schooling: while in college, individuals enroll in many courses and this fosters interaction among students who hone their communication skills and expand their social networks. Brown, Ivkovic, Smith, and Weisbenner (2008) find that stock holdings increase among long-time residents of a community when more stockholders move in, consistent with a learning or mimicking effect. van Rooij, Lusardi, and Alessie (2011) find that parents, friends, and acquaintances are the most common (self-reported) source of financial advice while professional financial advisors are the second most important source. More educated and wealthier individuals are more likely to consult professional advisors.

A separate channel, which we are not able to address in this paper, is trust. Individuals who have less trust in "the stock market"—firm owners, executives, brokers or fund managers, etc.—invest less in stocks as documented by Guiso, Sapienza, and Zingales (2008). Lack of trust may be more common among those who do not understand the workings of the stock market and the legal protection of investors. Guiso, Sapienza, and Zingales (2008) find that both college education and trust are significant in most of their specifications but they do not provide evidence that education has an impact on trust.

More risk averse agents will participate less in the stock market in the face of monetary entry costs under typical assumptions about returns and utility functions. Risk aversion is often hard to measure, which may be why Guiso, Sapienza, and Zingales (2008) find evidence of such a pattern while van Rooij, Lusardi, and Alessie (2011) do not. Dohmen, Falk, Huffman, and Sunde (2010) demonstrate, using a random sample of Germans, that risk aversion is a function of cognitive ability, while Hryshko, Luengo-Prado, and Sørensen (2011), using compulsory schooling laws as instruments, find that parental

education, more so than own education, affects the likelihood of being highly risk averse.

Ameriks, Caplin, and Leahy (2003) show that households with higher propensities to plan—in particular, those with better mathematical skills and those who are keen on detailed vacation planning—accumulate more wealth, which may lead to higher college enrollment of their children. In the face of fixed monetary costs of participation, individuals with higher wealth or income will have a higher propensity to hold stock. Wealth and income are clearly endogenous to education and likely also to affect risk aversion, trust, IQ, education, etc.; however, it is still informative to check if the results are robust to the inclusion of wealth and/or income: if, say, the effect of education is robust to the inclusion of wealth and income then the effect of education does not solely work through higher wealth and income of educated individuals. Lovenheim (2011) provides direct evidence on the importance of wealth for education by showing that, arguably exogenous, house appreciation was associated with a higher propensity of families to send their children to college during the housing boom of the 2000s. A corollary is that the magnitude of college tuition and fees matter.

3 PSID results

3.1 The PSID data

We use data from the PSID panel which started in 1968 interviewing about 4,800 households and follows individuals and their descendants over time. The PSID has conducted annual interviews collecting extensive socioeconomic information, with a switch to a biennial frequency from 1997 onwards.

3.1.1 The measure of stock ownership

We compile household wealth, income, and demographic data from the PSID. Information on household wealth is obtained from the PSID wealth supplements available at five-year intervals starting in 1984, and every other year from 1999 to 2007. Households may hold stock directly, by purchasing shares in publicly traded companies and/or mutual funds, or indirectly, in their pension funds and retirement accounts. In 1984 and 1989, heads of household were asked the following question about family stockholding:

> "Do you (or anyone in your family living there) have any shares of stock in publicly held corporations, mutual funds, or investment

trusts, including stocks in IRAs?"

There are some inconsistencies in how the PSID collects information on stock market participation over time. In 1999, 2001, 2003, 2005, and 2007 the PSID asked respondents about stock holdings, disregarding holdings in employer-based pensions or IRAs and, in those years, there is a separate question on whether the household had any money in private annuities or IRAs. In 1994, the question about stock ownership excluded stocks held in IRAs and employment-based pensions but the PSID did not include a separate question about money held in IRAs and private annuities. Thus, for the 1994 through 2007 wealth supplements, we are able to construct a measure of households' stockholding outside of retirement accounts while we cannot separate out stocks in retirement accounts when using 1984 and 1989 data. Our measure of stock ownership is a dummy variable equal to one if a household holds stocks outside of retirement accounts, zero otherwise. For brevity, we will use the shorthand "non-retirement stockholding" for holdings of stocks and mutual funds outside of retirement accounts. Our results are similar if stock holdings in retirement accounts are included in stock ownership.[5]

Table 1 provides some summary statistics for the PSID data used in our empirical analysis. The share of households that own stock outside of retirement accounts is stable over time at about 25% (although the total fraction owning stock, outside or inside retirement accounts, has increased steadily over time).

3.1.2 Construction of the environmental variables and the instrument for college graduation

Our sample selection rules are the following. First, we drop individuals who are not heads of household in any of the years when household wealth is recorded. We then construct consistent measures of the head's race, age, and background variables—the details are provided in Appendix A.

Individual schooling and holdings of risky assets may be affected by the quality of the environment an individual experiences in childhood and adolescence. We are able to construct some measures of the overall "quality" of the county in which the respondent grew up.[6] We focus on county median income but also explored the effect of the percentage of urban population and

[5]See Table 5, column (6) and the discussion in Section 3.2.2.

[6]We obtain county-level information from Haines (2004) who compiled county-level data for 1790–2000 from historical decennial censuses and, for the more recent years, county data books.

the median house value in the county (results not reported for brevity).[7] In addition, the PSID allows us to measure individual-specific background variables. In particular, heads of household were asked about their fathers' and mothers' schooling, whether they lived with both parents, and whether their family was poor, of average well-being, or rich when the head was growing up.

Because father's and mother's schooling are highly correlated and our sample sizes are not large, we cannot statistically distinguish their separate effects on individual schooling or holdings of risky assets and we combine the information into a single parental education variable set to the maximum of the mother's and father's education dummies. We call this variable parents' education. It is equal to zero if neither parent completed high school and equal to one if at least one of the parents finished high school. The typical head of household in our sample grew up in a family with at least one parent who finished high school (see Table 1).

We define a "Rich Parents" dummy which equals one if the head reports his or her family was rich and equal to zero otherwise.[8] We also construct consistent measures of whether the head lived with both parents while growing up and whether he or she grew up on a farm or in a city. About 28% of heads in our sample recall growing up in a rich family and 16% grew up on a farm. The majority of our sample heads, about 73%, grew up in a family with two parents.

We utilize data on head's years of schooling from the individual file of the PSID. Education records, first collected in 1968, are not updated annually; rather, education is first recorded at the time an individual enters the PSID as head of household and it is updated if an individual reenters the survey as head after an interrupted headship. It was also updated for any head, new or existing, in 1975 and 1985. Our samples consist of heads who completed their education by the time they turned 23 and have consistent education records.

In the 1970–1993 waves, the PSID collected information on the county and state where the respondent grew up. Because we do not have data on the actual county of residence at age 17 (the age when college availability has potentially the largest effect on individuals' schooling decisions as they approach high school graduation), we utilize this variable to assign respondents to counties. We drop heads with inconsistent records on the state and/or county of growing

[7]These measures are highly correlated and we decided to control just for median income in the county when an individual was 17 in all the regressions. Our results are robust to inclusion of all three county measures.

[8]Since the PSID follows split-offs of original families over time, in principle, it is possible to collect actual parental wealth for some individuals. However, samples constructed this way are too small for reliable regression analysis and we use recall variables instead.

Table 1: Summary statistics. Panel Study of Income Dynamics Data

Variable	Mean	Std. Dev.	Min.	Max.	N
Total stockholding					
Ever owned stock	0.414	0.493	0	1	4840
Owned stock 1984	0.172	0.378	0	1	2602
Owned stock 1989	0.211	0.408	0	1	3096
Owned stock 1999	0.396	0.489	0	1	2339
Owned stock 2001	0.428	0.495	0	1	2297
Owned stock 2003	0.420	0.494	0	1	2225
Owned stock 2005	0.420	0.494	0	1	2110
Owned stock 2007	0.429	0.495	0	1	2030
Non-retirement stockholding only					
Owned stock 1994†	0.268	0.443	0	1	3634
Owned stock 1999	0.247	0.431	0	1	2339
Owned stock 2001	0.271	0.445	0	1	2297
Owned stock 2003	0.246	0.431	0	1	2225
Owned stock 2005	0.245	0.43	0	1	2110
Owned stock 2007	0.245	0.43	0	1	2030
Education (years)	12.814	2.054	4	16	4840
Parents' education, dummy	0.699	0.459	0	1	4740
Lived with both parents	0.73	0.444	0	1	4615
Rich Parents	0.283	0.451	0	1	4489
Year of birth	1957	7.378	1943	1976	4840
County Median Inc./10000	2.506	0.681	0.473	5.336	4791
Log average income	9.878	0.854	5.006	12.888	4834
Log average net worth	7.793	5.65	–12.989	16.046	4840
Log average wealth w.o. stock	7.692	5.674	–12.989	15.993	4840
Family size	2.891	1.288	1	10.333	4840
Number of children	1.087	1.002	0	6.25	4840
College 4/1000	0.088	0.155	0	4.747	4806
College 4/1000 (winsorized)	0.084	0.106	0	0.553	4806
College 2/1000	0.055	0.101	0	2.849	4806
College 2/1000 (winsorized)	0.053	0.08	0	0.396	4806
Grew up on a farm	0.155	0.362	0	1	4746
White	0.596	0.491	0	1	4840
Married	0.639	0.48	0	1	4839
Male	0.743	0.437	0	1	4840
<=12 years of schooling	0.617	0.486	0	1	4840
13–15 years of schooling	0.188	0.39	0	1	4840
College graduate	0.195	0.396	0	1	4840

Notes: "Non-retirement stockholding" refers to ownership of stocks and mutual funds outside of retirement accounts while "Total stockholding" adds stocks and mutual funds in retirement accounts. †In 1994, there is no separate question on stock holdings in retirement accounts in the PSID, so we report sample statistics on non-retirement stockholding only in that year.

up.[9]

The instrument for own schooling is the number of colleges per 1,000 college-age persons in the county where the head grew up in the year when the respondent was 17 (college-age defined as being 18–22 years of age). Currie and Moretti (2003) construct a dataset that contains the availability of colleges in U.S. counties for 1960–1996. Our final sample contains only heads who turned 17 during this period.

The average head turned 17 in 1972 and grew up in a county with 0.09 colleges per 1,000 college-age persons (see Table 1). There is substantial variation in the availability of colleges, with "college-scarce" counties having zero colleges, and "college-abundant" counties having nearly 5 colleges per 1,000 persons aged 18–22. The distribution of colleges across counties displays several large outliers, reflecting clusters of colleges in small counties and we winsorize this variable at the 99% tail to avoid undue influence from these counties. Household heads, on average, are high school graduates, are predominantly male (74%), and 64% have been married at least once during the sample period.

3.2 Estimation using PSID data

3.2.1 The effect of local college density on education

Currie and Moretti (2003) study the effect of maternal education on health outcomes of children at birth. We follow Currie and Moretti (2003), instrumenting years of schooling with the number of colleges per 1,000 college-age persons in the head's county when he or she was 17, where "head's county" is short-hand for the county in which the head grew up. Currie and Moretti (2003) provide a detailed discussion of the validity of the instrument. For our purposes, the instrument is valid if it is effective in predicting education and it is unrelated to unobservable variables that affect household risky asset holdings such as, for example, heads' or their parents' attitudes towards risk or household earning capacities. We present some evidence on the effectiveness of the instrument in Tables 2 and 3. The most likely source of correlation between parents' attitudes and college availability would be if certain parents systematically moved to counties with more colleges. Currie and Moretti (2003) explore this issue in detail and find little evidence of such a pattern. They further guard against such correlations by including county dummies. Our sample is too small for doing this but we include regional dummies for

[9]Multiple records on the state and/or county where the head grew up are possible if an individual reenters the PSID as head after an interrupted headship.

the Census division where the respondent grew up.[10] Another potential problem is that colleges are not randomly assigned to counties. Universities may be opened in wealthier areas where parents can afford to pay tuition, and one might worry that the number of 4-year colleges in a county correlates with features of the county such as high average wealth and education that might lead to children becoming more financially savvy. Our county-level income variable helps alleviate this potential problem.

In Table 2, we (cross-sectionally) regress individual years of schooling on the number of 4-year colleges per 1,000 persons aged 18–22 in the heads' county at age 17. In addition, we control for parental education, an indicator for growing up in a rich family, an indicator for growing up with both parents, and median income in the county when the head was 17. More recent cohorts attain, on average, more years of schooling and have access to more colleges in their county. We therefore include year dummies together with age in order to control for the correlation between the availability of colleges and individual years of schooling due to aggregate trends in schooling.[11] Geographical areas may have different endowments and industrial structure (e.g., agricultural versus manufacturing states) and therefore may permanently differ in their demand for an educated workforce. Areas with relatively higher demand for skilled workers may attract more educated individuals (parents) and build more colleges in order to support a sustainable supply of skilled workers. If parental education is higher in some regions and is correlated with unobserved traits that affect offspring's education, we might find the number of colleges correlating with education even if there were no direct causal effect on education from college availability. To hedge against such effects, we include dummies for the region where the head grew up.

In column (1), we present results for the entire sample which contains 4,094 individual observations. We find that adding one more 4-year college per thousand college-age persons, holding everything else constant, increases individual education by about 0.6 years. One college more per thousand is, of course, quite a lot, so we interpret the coefficient as 0.06 more years of education on average if one college is added per 10,000 college-age persons (which is about one standard deviation of the variation of colleges in our sample). Assuming that it takes 4 years to obtain a college degree, this translates into 150 more college graduations per 10,000.[12] The effect is nearly significant at the ten per-

[10]Currie and Moretti (2003) verify, for a smaller sample from the National Longitudinal Survey of Youth, that the first stage estimates are very close whether county or only state dummies are included.

[11]Age and year dummies refer to the year of the first record on head's education.

[12]150 is obtained as $0.06 \times 10{,}000/4$; this number closely corresponds to our estimates in

Table 2: Regressions of Own Education on College Availability

	All		Poor or Avg. (All Races)		Poor or Avg. White	Poor or Avg. Non White
	(1)	(2)	(3)	(4)	(5)	(6)
College 4/1000	0.564	0.568	0.772*	0.768*	1.020**	–0.112
	(1.61)	(1.63)	(1.87)	(1.85)	(2.32)	(–0.13)
College 2/1000		0.807*		0.768	–0.004	2.263***
		(1.69)		(1.41)	(–0.01)	(2.94)
White	0.420***	0.412***	0.270***	0.261***		
	(4.91)	(4.83)	(2.74)	(2.65)		
Male	–0.047	–0.050	–0.007	–0.009	–0.092	0.081
	(–0.65)	(–0.69)	(–0.09)	(–0.11)	(–0.76)	(0.65)
Parents' edu: HS sum	1.443***	1.447***	1.441***	1.444***	1.791***	1.003***
	(18.09)	(18.00)	(17.64)	(17.56)	(14.53)	(9.12)
Lived with both parents	0.300***	0.303***	0.280***	0.284***	0.438***	0.077
	(4.06)	(4.09)	(3.31)	(3.35)	(3.57)	(0.65)
Rich Parents	0.194***	0.198***				
	(3.08)	(3.15)				
Grew up on a farm	–0.304***	–0.308***	–0.311***	–0.311***	–0.349***	–0.339*
	(–3.17)	(–3.20)	(–2.72)	(–2.74)	(–2.60)	(–1.69)
County Median Inc./10^3	0.234***	0.247***	0.238***	0.251***	0.329***	0.159
	(3.28)	(3.47)	(3.02)	(3.20)	(3.55)	(1.22)
Age	0.283***	0.282***	0.253***	0.252***	0.276***	0.182**
	(6.75)	(6.71)	(5.35)	(5.33)	(4.81)	(2.54)
Age squared/100	–0.351***	–0.348***	–0.314***	–0.311***	–0.329***	–0.244**
	(–5.34)	(–5.27)	(–4.29)	(–4.24)	(–3.68)	(–2.22)
Constant	4.916**	4.761**	5.626***	5.460***	4.848***	6.464***
	(2.25)	(2.17)	(6.41)	(6.22)	(4.82)	(4.81)
Region grew up dumm.	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
Adj. R sq.	0.210	0.210	0.188	0.188	0.204	0.097
N	4094	4094	2933	2933	1776	1157

Notes: The left-hand side variable is the respondent's reported years of completed schooling. "College 2/1000" is the number of 2-year colleges per 1,000 persons aged 18–22 in the county where the respondent grew up when s/he was 17. "College 4/1000" is defined analogously for the number of 4-year colleges. Both college variables are winsorized at the 99% level. "Year" is the first year education is recorded and "age" is age in that year. Standard errors in parentheses clustered by the county where the respondent grew up. "Parents' edu: HS sum" is a dummy variable equal to 0 if parents did not finish high school (HS); 1—if one or both parents finished HS (or more). *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

cent level. Children who grew up in better environments—in more educated, richer, and stable families, and in high-income counties—attain higher levels of schooling.

In column (2), we add 2-year colleges—this variable is statistically significant at the ten percent level while the estimated impact of 4-year colleges is unchanged. In columns (3) and (4), we drop respondents who grew up in "rich" families who may not be constrained by the availability of local colleges. As expected, we find a stronger impact of 4-year colleges for this sample, but no significant effect of 2-year colleges.

Columns (5) and (6) examine non-rich white and non-white respondents, respectively. The effect of 4-year colleges on education is very large for non-rich white respondents. For non-white respondents, the effect of 2-year colleges is statistically significant with a large point estimate (0.226 more years of education for one more college per 10,000 college-age persons).

We want to ascertain that college availability affects college graduation by increasing the number of individuals who continue past high school. In Table 3, we first consider the effect of 4-year colleges on whether the head is a high school drop-out or finished high school and no more—column (1). One more 4-year college per 10,000 college-age persons in the county reduces the probability that the head finished 12 or less years of schooling by 1.5 percentage points. Similarly to Currie and Moretti (2003), we find that 4-year colleges do not affect "some college"—column (2), while the effect of 4-year colleges on the likelihood of individual college graduation is strong and significant at the one percent level for the sample of white heads from less privileged backgrounds. One more 4-year college per 10,000 college-age persons increases the probability of college graduation by about 1.2%, corresponding to 120 more graduates per 10,000. Thus, it appears that the increase in college graduations is due to a reduction in the number of persons who finish high school but do not take any college courses. The magnitude of the effect of 4-year colleges on the likelihood of graduating from college is larger for the sample of households whose heads grew up in less well-off families. For this subsample, the net effect is a decline in the number of individuals who stop education after graduating from high school, an unchanged number of individuals with some college, and an increase in the number of college graduates.

Table 3, column (3). The effect is somewhat smaller than the one found in Currie and Moretti (2003) but their result is based on a sample of females, while our sample consists predominantly of males as seen in Table 1. Perhaps, the difference is due to the fact that education of males is less affected by changes in the availability of local colleges.

Table 3: Regressions of Education Dummies on College Availability

	All			Poor-Average White		
	$\leq$ 12 yrs	13–15 yrs.	$\geq$ 16 yrs	$\leq$ 12 yrs	13–15 yrs.	$\geq$ 16 yrs
	(1)	(2)	(3)	(4)	(5)	(6)
College 4/1000	–0.150*	0.029	0.121*	–0.290***	0.034	0.256***
	(–1.88)	(0.53)	(1.74)	(–2.75)	(0.51)	(3.01)
White	–0.089***	–0.027*	0.115***			
	(–4.67)	(–1.90)	(7.87)			
Male	0.008	0.003	–0.011	0.012	–0.004	–0.008
	(0.45)	(0.21)	(–0.80)	(0.37)	(–0.17)	(–0.30)
Parents' edu: HS sum	–0.287***	0.114***	0.173***	–0.354***	0.104***	0.250***
	(–17.82)	(9.41)	(12.37)	(–13.29)	(5.15)	(12.76)
Lived with both parents	–0.058***	0.006	0.051***	–0.057*	–0.002	0.059**
	(–3.12)	(0.46)	(3.80)	(–1.86)	(–0.09)	(2.28)
Rich Parents	–0.040**	0.002	0.038***			
	(–2.40)	(0.11)	(2.79)			
Grew up on farm	0.067***	–0.036**	–0.030*	0.063**	–0.026	–0.037
	(3.09)	(–2.17)	(–1.78)	(1.98)	(–1.15)	(–1.31)
County Median Inc./10^3	–0.051***	–0.000	0.052***	–0.083***	0.005	0.078***
	(–3.18)	(–0.02)	(3.49)	(–3.81)	(0.30)	(3.80)
Age	–0.058***	0.010	0.049***	–0.062***	0.005	0.056***
	(–6.13)	(1.25)	(6.61)	(–4.58)	(0.47)	(4.35)
Age squared/100	0.071***	–0.014	–0.057***	0.074***	–0.008	–0.066***
	(4.77)	(–1.22)	(–4.88)	(3.54)	(–0.47)	(–3.25)
Region grew up dumm.	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
Adj. R sq.	0.154	0.031	0.141	0.143	0.023	0.110
N	4094	4094	4094	1776	1776	1776

Notes: The left-hand side variable is a dummy equal to one if the respondent's years of schooling fall into any of the indicated categories. "College 4/1000" is the number of 4-year colleges per 1,000 persons aged 18–22 in the county where the respondent grew up when s/he was 17, winsorized at 99%. "Year" is the first year education is recorded and "age" is age in that year. Standard errors in parentheses clustered by the county where the respondent grew up. "Parents' edu: HS sum" is a dummy variable equal to 0 if parents did not finish high school; 1—if one of them finished HS or more. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

3.2.2 The effect of college graduation on equity holdings. Probit and IV-probit regressions

Table 4 presents probit regressions of non-retirement stock ownership on a dummy for college graduation and the estimates are presented in terms of the marginal impact (evaluated at the mean) on the probability of owning equity outside of retirement accounts.[13] Our probit estimations "regress" household risky equity ownership on heads' education, parents' education, and exogenous background variables and demographic controls. In these regressions, we use multiple observations per household head taking advantage of the panel nature of the PSID and we cluster standard errors by household head.[14] Consistent with previous studies, we find that more educated households have larger propensities to own equity. Holding other variables constant, graduating from college increases the probability of owning equity by 25.8%. Male-headed households and older households are more likely to own stock, as are households whose heads have more educated parents and those who grew up in high-wealth households or in high-income counties.

Own education is correlated with many unobserved household characteristics, such as preferences towards risk and abilities, and it may capture the effects of these omitted variables on the probability of non-retirement equity holding. It is hard to know how those left-out variables may bias the results although risk tolerance would likely be positively correlated with college graduation and if this is an important left-out variable one would expect upward bias in the OLS estimates. On the other hand, measurement error in education would lead to downward bias in OLS estimates.

To eliminate such effects, we use as an instrument the number of 4-year colleges when the head was 17, in the county where he or she grew up. Column (2) in Table 4 reports our results. The point estimate for the effect of education is larger than in the non-instrumented regression but not significant. The first-stage F-statistic for significance of the number of colleges in column (2) is 2.63. A value less than 10 is usually considered an indicator of weak instruments and the estimates in the second column should therefore be taken with a grain of salt.

We further focus on the effect of college graduation for the sample of house-

[13] In the literature many papers use simple OLS-regressions on cross-sectional data. Results from linear regressions were similar to the probit results and we do not display them.

[14] The discrepancy between the number of heads, say, in the sample of column (1) of Table 3 and column (1) of Table 4 is that in Table 3 individuals who have observations on stock/IRA holdings only in 1984 and/or 1989 are included but dropped in Table 4 as we focus is on non-retirement stockholding which can be cleanly measured only since 1994.

Table 4: Panel Probit Regressions of Risky Assets' Ownership on College Dummy: Non-Retirement Stockholding Only

	Probit	IV-Probit	IV-Probit	
	Total sample	Total sample	Poor or Avg.	White
	(1)	(2)	(3)	(4)
Education dummy: 16 or more yrs.	0.258***	0.323	0.497**	0.476**
	(14.36)	(0.60)	(2.20)	(2.00)
Age	0.026***	0.026***	0.026***	0.019**
	(5.63)	(3.69)	(2.69)	(2.41)
Age sq./100	–0.020***	–0.020***	–0.021**	–0.018*
	(–3.63)	(–3.59)	(–2.40)	(–1.83)
White	0.167***	0.160**		
	(11.23)	(2.49)		
Male	0.115***	0.114***	0.117***	0.005
	(7.52)	(6.88)	(3.05)	(0.11)
Parents' edu (dummy)	0.074***	0.064	0.055	0.033
	(4.53)	(0.81)	(0.66)	(0.46)
Lived with both parents	0.009	0.006	0.006	–0.012
	(0.53)	(0.21)	(0.18)	(–0.37)
Rich parents	0.041***	0.038		
	(2.77)	(1.41)		
Grew up on farm	–0.013	–0.011	–0.002	–0.013
	(–0.71)	(–0.40)	(–0.06)	(–0.41)
County median inc./100	0.050***	0.046	0.030	0.023
	(3.80)	(1.22)	(0.77)	(0.66)
Married				0.055*
				(1.80)
Number of children				–0.026***
				(–2.79)
Log income				0.038
				(0.99)
Log wealth				0.021***
				(4.48)
Year dummies	Y	Y	Y	Y
Region grew up dummies	Y	Y	Y	Y
F-Instrument		2.63	7.61	7.96
χ^2	967.4	760.9	323.5	490.2
N clusters	3269	3269	1457	1457
N	12754	12754	5838	5838

Notes: The left-hand side variable is a dummy equal to 1 if a household owns stock, 0—otherwise. Marginal effects evaluated at the mean of the explanatory variables are reported. The instrument for years of schooling is "College 4/1000" winsorized at the 99 percentile. t-statistics in parentheses. Standard errors clustered by individual. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

holds with white heads born in non-rich families since the combined results of Tables 2 and 3 point to the robust importance of colleges for stockholding of this segment of the U.S. population. In columns (3)–(4) of Table 4, the results are similar to those of the previous two columns but significantly stronger and, in particular, the F-statistic for the instrument is much larger, around 8, in spite of the smaller sample, in line with the finding that the number of colleges has its main effect on non-rich whites.[15]

In column (4), we include some further, potentially endogenous, controls (income, wealth, marital status, and family size).[16] Clearly, income and wealth are endogenous to education and wealthier households have larger propensities to own stock—the wealth coefficient is significant at the one percent level—consistent with previous studies (e.g., Campbell 2006, Vissing-Jørgensen 2002b). We find that family size is inversely related to equity ownership, maybe due to households with more children being more risk averse or needing liquid savings more. The inclusion of income and wealth does not change the estimated impact of education on stockholding much for white heads with non-rich parents and the coefficient is significant at the five percent level whether these variables are included or not.

It is somewhat puzzling that the IV-estimates are much larger than the OLS-estimates. However, this pattern has been found in related studies using similar instruments to estimate the effect of schooling on labor market outcomes. Card (2001), in an important paper, provides the following summary: "One interpretation of this finding is that marginal returns to education among the low education sub-groups, typically affected by supply-side innovations tend to be relatively high, reflecting their high marginal costs of schooling, rather than low ability that limits their return to education." We subscribe to this interpretation and conjecture that our instrument mainly affects the disadvantaged. Individuals face a trade-off between the cost of college and the benefits of having a college degree. If a college gets built nearby, the cost of attending college for disadvantaged individuals goes down as students can live at home instead of having to move and pay dorm fees. Prospective students with less wealthy parents are more likely to be credit constrained and

[15]Stock, Wright, and Yogo (2002) suggest (in their Table 1) that an F-statistic of 8.96 guarantees that the standard critical value of 1.96 for the t-statistic has a size of at most 15%. Our setup is a little more complicated than that of Stock, Wright, and Yogo (2002) (involving other regressors) but the reported t-statistic is probably best interpreted as signifying significance at around the 10-15% level.

[16]Our measure of wealth is household net worth inclusive of net business wealth; income is the average combined labor and transfer income of the head and wife for 1981–1997, 1999, 2001, 2003, 2005, and 2007.

the higher availability of colleges affects these individuals disproportionately. At the same time, individuals from less advantaged families are less likely to have acquired financial literacy from their parents and the marginal effect of schooling is therefore higher for this group.[17] The combination of college availability affecting schooling more and schooling affecting stockholding more for the less advantaged is what can explain the higher IV-coefficient—this pattern has been found in many other contexts involving instruments for schooling as spelled out by Card (2001).[18] If one accepts this argument, an implication is that one needs to be careful interpreting the estimated coefficients, e.g. for policy purposes—opening further colleges in well-off counties may have no effect while there may be a large effect in less well-off counties.

The large IV-coefficient to college graduation will mechanically depress the coefficients to variables such as parental education which are correlated with the instrument. Likely, parental education is an important determinant of college and stockholding, but the goal in this section is not to map out the role of parental education but rather to establish the causal effect of education on stockholding.

Overall, the effect of education, moved by changes in the number of local colleges, on individuals from poor or average families is quite substantial. Our results suggest that the construction of colleges in college-poor counties may be an effective policy not only for increasing the number of skilled workers but also for increasing savings of individuals from less advantaged backgrounds. Since risky equity, on average, generates higher wealth, increasing the education levels of the disadvantaged may better prepare them for retirement and help them buffer adverse shocks to income.

Robustness

In Table 5, we report a series of IV-probit regressions exploring the robustness of our results. In column (1), for convenience, we reproduce our results of Table 4, column (4). More risk averse individuals may attain more schooling and select less risky occupations. College education may therefore proxy for lower background (labor income) risk which directly causes higher incidence of stockholding. In column (2), we control for the full set of one-digit occupation dummies, constructed from information about the head's first job.

[17] We do not find significant effects using the highly disadvantaged Survey of Economic Opportunity (SEO) sub-sample of the PSID. Probably individuals in this group are unlikely to ever own stock whether a college gets opened nearby or not.

[18] Hryshko, Luengo-Prado, and Sørensen (2011) find the same pattern for the relations between compulsory schooling laws, education, and children's risk aversion and spell out the argument in more detail in a simple setting.

Our main result for the effect of college education on stockholding remains unaltered. In column (3), we control for another measure of background risk constructed from residuals of cross-sectional regressions of log-labor income of the head and wife on a third-order polynomial in head's age. Since the PSID changed to a biennial sampling frequency in 1997, we focus on the variance of two-year changes in residuals—a measure of the variance of idiosyncratic shocks to household income during the two-year span.[19] The variance measure is not significant and controlling for background risk does not change our main results.

Household wealth may affect stockholding non-linearly. In column (4), we add the square of household wealth to the set of controls. This results in an insignificant linear wealth term, while the effect of college education on stockholding retains its significance at the five percent level and we therefore proceed using the simpler linear term. Wealth may correlate with the likelihood of stockholding as those participating in the stock market accumulate larger wealth on average. In column (5), we include wealth, net of stocks, resulting in an estimated effect of wealth of similar magnitude but less precisely estimated than when using total wealth, while the estimated effect of college graduation is largely unchanged.[20]

In column (6), we consider a measure of stockholding that includes stocks in retirement accounts. The effect of college graduation is of similar magnitude and statistical precision as our previous estimates. Limiting our sample to a cross-section of households, where stockholding is defined as a dummy which equals one if a household ever owns stock, does not change our results—see column (7).[21] In columns (1)–(6), we cluster standard errors by individual as we typically have more than one observation for each individual. In the cross-sectional regressions of columns (7)–(8), we use White robust standard errors and standard errors clustered by county of growing up, respectively, to verify robustness of the estimated standard errors. In the presence of labor income risk and costly adjustment of homeownership, committed mortgage

[19]To make the measure more precise, we limit our estimation sample to the households with at least three observations on idiosyncratic two-year growth rates.

[20]In an analogous regression with household net worth net of both stocks and business wealth, the coefficient on the wealth term is 0.013, slightly more precisely estimated (with a t-statistics of 4.12), while the estimated coefficients on all other regressors and their precision are virtually unchanged from column (4). These results are not tabulated for brevity.

[21]In these regressions, head's age is measured at the year of the first record on household stockholding, the "Married" dummy equals one if a head of household is ever married during the sample period, the number of children and wealth is measured as the average number of children and wealth, and income is measured as the average of the combined head's and wife's labor and transfer income during the sample period.

Table 5: IV-Probit Regressions of Non-Retirement Stock Ownership on College Dummy: Sample of Whites from Non-Rich Families. Robustness

	(1) Original	(2) Occupation dummies	(3) Variance Inc. Growth	(4) Wealth Squared	(5) Wealth No Stock$^{\triangle}$	(6) Direct & Indirect	(7) Cross Section	(8) Cross Section$^{\sharp}$	(9) Home Owners	(10) Home Renters	(11) Parent No Stock$^{\diamond}$
College dummy	0.476**	0.589***	0.454*	0.479**	0.486**	0.438**	0.534**	0.534**	0.372	0.736***	0.723***
	(2.00)	(2.68)	(1.82)	(1.98)	(2.08)	(1.98)	(2.06)	(1.96)	(1.27)	(3.62)	(3.39)
Age	0.019**	0.017**	0.019**	0.016**	0.021***	0.002	0.008	0.008	0.022**	0.010	0.007
	(2.41)	(2.10)	(2.32)	(2.04)	(2.63)	(0.22)	(0.44)	(0.43)	(2.01)	(0.90)	(0.25)
Age sq./100	–0.018*	–0.016*	–0.018*	–0.020**	–0.019**	0.007	–0.008	–0.008	–0.020*	–0.016	–0.018
	(–1.83)	(–1.75)	(–1.80)	(–2.13)	(–1.98)	(0.58)	(–0.34)	(–0.34)	(–1.67)	(–0.97)	(–0.44)
Male	0.005	0.037	0.009	–0.002	0.010	–0.042	–0.059	–0.059	–0.015	0.050	–0.046
	(0.11)	(0.81)	(0.19)	(–0.04)	(0.23)	(–0.81)	(–0.57)	(–0.52)	(–0.24)	(1.00)	(–0.45)
Parents' edu (dummy)	0.033	–0.000	0.039	0.002	0.034	0.072	–0.012	–0.012	0.057	–0.030	0.003
	(0.46)	(–0.00)	(0.53)	(0.04)	(0.47)	(0.88)	(–0.14)	(–0.13)	(0.71)	(–0.24)	(0.02)
Lived with both parents	–0.012	–0.001	–0.014	–0.021	–0.007	–0.013	–0.000	–0.000	–0.009	–0.045	–0.064
	(–0.37)	(–0.03)	(–0.42)	(–0.67)	(–0.21)	(–0.38)	(–0.01)	(–0.01)	(–0.24)	(–0.96)	(–0.94)
Grew up on farm	–0.013	–0.009	–0.014	–0.021	–0.011	0.015	–0.043	–0.043	–0.022	0.037	–0.035
	(–0.41)	(–0.29)	(–0.44)	(–0.64)	(–0.34)	(0.46)	(–1.04)	(–0.99)	(–0.63)	(0.66)	(–0.46)
County median inc./100	0.023	0.014	0.027	0.006	0.023	0.081*	–0.000	–0.000	0.050	–0.070	–0.050
	(0.66)	(0.44)	(0.75)	(0.18)	(0.65)	(1.90)	(–0.01)	(–0.01)	(1.44)	(–0.99)	(–1.05)
Married	0.055*	0.031	0.051	0.025	0.062**	0.131***	0.149***	0.149***	0.031	0.026	0.137
	(1.80)	(0.94)	(1.59)	(0.73)	(2.02)	(3.84)	(2.94)	(2.78)	(0.72)	(0.65)	(1.17)
Number of children	–0.026***	–0.022**	–0.027***	–0.025***	–0.026***	–0.030***	–0.028	–0.028	–0.033***	–0.001	–0.005
	(–2.79)	(–2.28)	(–2.78)	(–2.69)	(–2.71)	(–2.83)	(–1.31)	(–1.45)	(–2.70)	(–0.04)	(–0.14)
Log income	0.038	0.029	0.045	0.013	0.042	0.055	0.220	0.220	0.036	0.059	0.128
	(0.99)	(0.86)	(1.09)	(0.43)	(1.06)	(1.29)	(1.00)	(0.95)	(0.78)	(0.92)	(0.30)
Log wealth	0.021***	0.020***	0.022***	0.003	0.015***	0.026***	0.022***	0.022***	0.037***	0.006**	0.014***
	(4.48)	(3.72)	(4.47)	(1.29)	(3.93)	(5.26)	(3.07)	(2.86)	(3.34)	(2.56)	(2.12)
Log wealth sq/100				0.438***							
				(3.07)							
Variance log inc. growth			0.006								
			(0.83)								
Occupation Dummies	N	Y	N	N	N	N	N	N	N	N	N
Year Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Region Grew Up Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
F-Instrument	7.96	6.77	7.91	7.60	7.99	7.89	5.95	5.69	6.64	2.54	1.23
χ^2	490.2	624.0	462.9	877.8	508.8	741.1	527.8	535.6	291.8	449.4	287.3
N clusters	1457	1363	1347	1457	1457	1457	1461	122	1161	549	425
N	5838	5438	5717	5838	5838	5838	1461	1461	4557	1127	425

Notes: See notes to Table 4. t-statistics in parentheses. Standard errors clustered by individual unless an exception is noted. "Cross-section" is cross-sectional with the left hand side equal to one if the respondent ever report holding stock. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. $^{\triangle}$: Wealth measures excludes value of stocks. $^{\sharp}$: Clustering by county where the respondent grew up. $^{\diamond}$: Respondent and parent matched data.

payments may make owners more reluctant to expose themselves to stock-market risk. Yao and Zhang (2005) show that homeowners optimally hold less stock and provide evidence from the PSID that the probability of stock market participation increases more steeply in income and wealth for renters compared to owners. We find, see columns (9)–(10), a similar result for income (although this result is not significant) and we find that college education affects the probability of holding stock more for renters—however, the effect of college for owners is not precisely estimated and the difference between owners and renters is not statistically significant; also, the first-stage F-statistic is rather small for renters.[22]

Finally, parental stockholding may be an important left-out variable—this could potentially bias our results if parents who hold stocks are more likely to move to college-rich areas at the same time as parental stockholding causes children's stockholding. In column (11), we use matched parent-children data and show that our results are not driven by stockholding parents moving to counties with more colleges because the IV-coefficient to children's education is, if anything, larger and more significant when restricting the sample to parents not owning stocks. This result, however, should be taken with caution as the first stage F-statistic indicates that our instrument is ineffective in predicting college graduation in the matched sample.

4 Channels (Estimations using WLS data)

Having established a causal effect of college education on non-retirement stockholding, we next provide a broad picture of the possible channels through which the effect manifests itself. We use data from the Wisconsin Longitudinal Survey (WLS), a long-term panel survey of a random sample of 10,317 men and women who graduated from Wisconsin's high schools in 1957. Although the WLS is limited in its geographical and cohort scope (rendering the college-density instrument ineffective for this sample), it is very rich in other dimensions which allows us to take a deeper look at channels than is possible with PSID data. We do not have instruments for the estimations in this section which is better seen as narrowing down potential explanations rather than finding the exact channel.

We explore different ways in which education may affect financial behavior including the effect of education on cognitive ability, occupational choice, wealth accumulation, marital status, and financial literacy associated with

[22]Our sample is different from that of Yao and Zhang (2005) because we focus on non-rich white households.

choice of major. Our strategy is to add these variables one-by-one to the (probability) regression of stockholding on a college dummy and examine which variables reduce the estimated impact of education. This approach is slightly unorthodox, in the sense that such a pattern in standard multiple regression analysis may simply imply left-out variable bias in the initial regression.

For example, cognitive ability after college is a potential channel for education to influence financial behavior. Including this variable lowers the coefficient of college, because cognitive ability and college education are positively correlated, indicating that college affects stockholding by increasing cognitive ability. What validates this interpretation is the result of the previous section that exogenous increases in college graduation affect equity holdings and this exogenous effect cannot be rendered insignificant by a component of cognitive ability which is not caused by education (except by chance). Of course, it is essential for our argument that cognitive ability is measured *after* college and the case is stronger if IQ before college is controlled for.

Consider first the case when college completion, $Coll_i$, and cognitive ability after college, Cog_i, share the influence of a common unobserved component, u_i. College completion is further affected by the number of colleges $NColl_i$ (as in our first-stage regression), an unobserved component v_i, and cognitive ability is affected by an unobserved component ϵ_i. Assume that v_i and ϵ_i are uncorrelated. These relations can be described by the following two equations: $Coll_i = \alpha_0 + \alpha_1 u_i + \alpha_2 NColl_i + v_i$, and $Cog_i = \beta_0 + \beta_1 u_i + \epsilon_i$. In this case, by the Frisch-Waugh theorem, the coefficient to college completion in an OLS regression of the incidence of stock ownership on college completion and cognition after college will largely depend on the covariation of stock ownership with $Coll_i$ and v_i.[23] From the first stage IV-regressions, we know that the number of colleges affects the probability of holding equity.[24] The number of colleges is exogenous and not a function of u_i and, therefore, the effect of college (via the number of colleges) on stock holdings cannot be rendered insignificant (except by chance) because of cognition capturing the effect of u_i. However, the

[23]To be precise, the coefficient to $Coll_i$ can be found by regressing the incidence of stockholding on the residual from regressing $Coll_i$ on Cog_i. For simplicity, normalize β_0 to zero; the residual, for the case outlined, equals $const + \alpha_1(1-\kappa)u_i + \alpha_2 NColl_i + v_i - \frac{\alpha_1}{\beta_1}\kappa\epsilon_i$, where $\kappa \equiv 1/(1 + \frac{\sigma_\epsilon^2}{\beta_1^2\sigma_u^2})$. If the variation in u_i is high and dominates the variation in cognition (so that κ approaches one), the inclusion of cognition will remove most of the variation in u_i from the college completion variable. Alternatively, if the variation in u_i is low, there is no left-out variable problem—in the OLS setting, the inclusion of cognition will not change the estimated coefficient to the college completion variable by much.

[24]In the linear probability model, our IV coefficient, ignoring the covariates other than college, equals $\frac{cov(P_i, NColl_i)}{cov(Coll_i, NColl_i)}$, where P_i denotes the incidence of stockholding.

effect of college availability on stockholding can be rendered insignificant by the inclusion of cognition after college in the probability estimate for stockholding if cognition itself is caused by college (and therefore by the number of colleges); i.e., $Cog_i = \tilde{\beta}_0 + \tilde{\beta}_1 u_i + \tilde{\beta}_2 NColl_i + \tilde{\epsilon}_i$. Our interpretation of the results is that the number of colleges in the county affects college attendance which affects cognitive abilities measured after college which again affects the probability of holding equity. This pattern is consistent with the IV- and OLS-estimates considered together.

Further, the WLS-data is well suited to control for unobserved variables that share influences on cognition, college attendance and stockholding, in particular through the inclusion of family fixed effects. The inclusion of such controls would make the "u"-term less important and make the causal effect from college to cognition more likely to be the dominant source of the correlation between these variables. The inclusion of other variables, besides cognition measured after college, in the regressions that follow is subject to similar considerations.

4.1 The WLS data

The WLS coverage is extensive with information on social background, schooling, labor market experience, family characteristics, social participation, psychological characteristics, etc. The survey has followed respondents throughout the life cycle, collecting data from the original respondents in 1957, 1964, 1975, 1992, and 2004. Most important for our study is the availability of information on asset holdings, as well as several cognition measures for each respondent (IQ-tests from high school and several cognition measures in the various survey years). The WLS has also collected information for a selected sibling of each original respondent (if not an only child) in 1977, 1994, and 2005, which allows us to control for unobserved family characteristics.

Our measure of stock market participation is constructed from the 2004 question (2005 for siblings) "Do you or your spouse have stocks, bonds, or shares in mutual funds?" We consider this to be a measure of "non-retirement" stock-bond market participation different from participation through retirement accounts because respondents are also asked "Do you or your spouse have any retirement plans that accumulate an account balance—examples include IRA's, 401(k) plans, and profit sharing plans." (Respondents also provide information on the value of these assets.) We focus on stock-bond holdings in 2004 and we use "current" (2004) controls except in the case of occupation

where we use prime-age data from 1992.[25]

Table 6 presents summary statistics for the key variables in the regression sample, which includes all main respondents who answered the stock-bond market participation question in 2004. 57% report non-retirement participation in financial markets. This figure is much larger than the fraction of PSID participants who report owning stock but this is not surprising because respondents in the WLS are on average more educated—they have at least finished high school. (Also, the WLS question regarding stock market participation is not exactly the same as the PSID question which does not include "bonds" in the wording.) The average number of years of schooling in the WLS sample is 13.7 with a standard deviation of 2 years, and 28% of the respondents have completed 4 or more years of college. To investigate the effect of major of study on stock market participation, we create a dummy variable that identifies respondents with an economics/business college degree, which is held by seven percent of the sample.[26] The average age for respondents in 2004 is 64 with a small standard deviation of 0.7 years and 48% of respondents are males. In 2004, 78% are married and have (ever had) three children on average.[27] There is a great deal of variation in net worth within the sample, with average and median wealth of $704,389 and $361,500, respectively.

To study the effect of occupation on stock-bond market participation, we construct a dummy variable that divides respondents into two groups according to their profession in the 1992 interview when respondents are on average 52 years old. We label a respondent "white collar" if he/she is a professional, a technical or kindred worker, a manager, an official, a proprietor, or a sales worker outside retail trade. Similarly, we construct an "aspiration white collar" dummy, which indicates if the respondent wanted to be in one of these professions when first interviewed in 1957. 50% of respondents are "white collar" in 1992 while 45% wanted to be "white collar" when they were finishing high school.

The WLS has several intelligence measures: a high-school IQ measure which accounts for differences in intellectual ability before entering college and a 2004 measure of cognition, arguably affected by college attendance. The high-school IQ score is mapped from a raw Henmon-Nelson test score (a

[25]In the 1992 survey, the questions regarding asset holdings are different and it is not possible to construct a clear measure of direct participation in financial markets. Respondents are asked if they have money on either savings or investment accounts.

[26]The majors included in our econ/business major dummy are accounting and finance, business, and economics and includes B.A., B.Sc., and graduate degrees in these fields.

[27]These include biological, adopted, step, or foster children as well as other children respondents considered to be part of their family.

Table 6: Summary statistics. Main Respondents. Wisconsin Longitudinal Data

Variable	Mean	Std. Dev.	Min.	Max.	N
Non-retirement stockholding	0.57	0.49	0	1	4741
College dummy	0.28	0.45	0	1	4741
Education (years)	13.66	2.29	12	20	4740
Econ/Business Major	0.07	0.26	0	1	4741
Age	64.29	0.67	63	67	4741
Male	0.48	0.5	0	1	4741
Married in 2004	0.78	0.41	0	1	4741
Number of children	3.03	1.7	0	10	4741
Log of number of children	1.29	0.49	0	2.4	4741
Wealth in 2004	704389.82	1253688.41	–15000	12000000	4741
Log of wealth in 2004	12.33	2.95	–9.62	16.3	4741
Total income 1992	67195.39	58365.56	0	520000	4741
Log of income in 1992	10.3	2.6	0	13.16	4741
White collar in 1992	0.5	0.5	0	1	4741
Wanted to be white collar	0.45	0.5	0	1	4741
IQ measure	103.12	14.32	61	145	4741
Log of IQ	4.63	0.14	4.11	4.98	4741
Cognition-Similarities 2004	6.74	2.33	0	12	4741
Log of cognition in 2004	1.99	0.35	0	2.56	4741
Father white collar	0.33	0.47	0	1	4741
Family well-off	0.22	0.41	0	1	4741
Parental income in 1957	43943.73	42065.17	672.24	670897.5	4741
Log parental inc. 1957	10.47	0.66	6.51	13.42	4741
Father education (years)	10.39	3.16	7	18	4741

Notes: "Non-retirement stockholding" refers to ownership of stocks, bonds or mutual funds outside retirement accounts.

30 minute test consisting of 90 items presented in order of increasing difficulty), which all Wisconsin secondary school students took from 1933 to the late 1950s or early 1960s according to the WLS documentation. The test includes "vocabulary, sentence completion, disarranged sentences, classification, logical selection, series completion, directions, analogies, anagrams, proverb interpretation, and arithmetic problems. Spatial, as well as verbal and numerical materials, are employed. The different types of items are not segregated but are arranged in a scrambled sequence." The average score for the respondents is 103 with a standard deviation of 14 points. Scores range from 61 to 145. The 2004 cognition measure is constructed from answers to nine of fourteen items from the Weschler Adult Intelligence Scale (WAIS). The WLS eliminates the five simplest items from the WAIS because "the general ability of the sample is high enough to cause little variation in response to simple items." For those not familiar with WAIS, typical questions would range from "in what way are an orange and a banana alike?" to "in what way are praise and punishment alike?" The scores range from 0 to 12 with an average of 6.7 and a standard deviation of 2.3.

In the regressions, we control for socioeconomic background with family fixed effects (for the sample of siblings) or by including the following variables (for the sample of main respondents): father's occupation (white collar or not), parental income in 1957, father's education, and a dummy variable that identifies respondents as being well-off when first interviewed in 1957.[28] 33% of respondents identify their parents as white collar, 22% consider themselves well-off when interviewed in 1957, fathers have on average 10.4 years of education, and average parental income is $43,940 in 1957 (2004 dollars).

Table 7 presents summary statistics for the relevant variables for the sample of siblings. Siblings' characteristics are quite similar, except siblings, by construction of the sample, have more variation in age (average age is 64 with a standard deviation of 7 years).

In several of the regression specifications, we use the logarithm of the variables described (income, wealth, number of children, intelligence measures, etc.). To make sure that we include observations where the variable in question takes the value of 0, we take the logarithm of the original variable plus 1. In the case of net worth, we use the transformation $\log(1+\text{abs}(\text{wealth}))\times \text{sign}(\text{wealth})$, to include respondents reporting negative wealth.[29]

[28]The exact wording of the question is "How does your family income or wealth compare with families in your community?" The dummy takes the value 1 if the respondent says "somewhat above average" or "considerably above average" and 0 otherwise (considerably below average, somewhat below average or average).

[29]We also apply this transformation when using PSID data.

Table 7: Summary statistics. Siblings. Wisconsin Longitudinal Data

Variable	Mean	Std. Dev.	Min.	Max.	N
Non-retirement stockholding	0.53	0.5	0	1	3972
College dummy	0.35	0.48	0	1	3972
Education (years)	13.95	2.59	0	21	3845
Econ/Business Major	0.04	0.19	0	1	3972
Age	63.8	7.13	34	87	3972
Male	0.48	0.5	0	1	3972
Married in 2004	0.76	0.43	0	1	3374
Number of children	2.9	1.81	0	10	3972
Log of number of children	1.24	0.52	0	2.4	3972
Wealth in 2004	582065.35	922561.89	–14508.45	6789954	3968
Log of wealth in 2004	11.73	3.86	–9.58	15.73	3968
Total income in 1992	60920.41	61022.11	–5000	500000	3297
Log of income in 1992	9.81	3.25	0	13.12	3294
White collar in 1992	0.49	0.5	0	1	3663
Wanted to be white collar	0.63	0.48	0	1	905
IQ measure	104.64	15.48	61	145	3292
Log of IQ	4.64	0.15	4.11	4.98	3292
Cognition-Similarities 2004	6.69	2.39	0	12	3693
Log of cognition in 2004	1.98	0.38	0	2.56	3693

Notes: "Non-retirement stockholding" refers to ownership of stocks, bonds or mutual funds outside retirement accounts.

4.2 Regression Results

We show results using a linear regression framework.[30] We focus on linear regressions instead of probability models because we want to include family fixed effects which is more naturally done in the linear regression framework. In the first column of Table 8, we find a coefficient of 0.159 to college education (i.e., a college education increases the probability of positive non-retirement stockholdings by about 16%).[31] The coefficient to college is smaller than the coefficient to college found for the PSID but this is likely due to the WLS sample having more education on average. Age and being a male are negatively significant, while high-school IQ is strongly positively significant, see column (2); the latter is consistent with the findings of Grinblatt, Keloharju, and Linnainmaai (2011).[32] Including IQ does not decrease the estimated effect of college by much so the estimated effect of college graduation is not just capturing higher IQ of college graduates. (We do not include age squared in the regressions reported in Table 8 because there is little age variation amongst the WLS main respondents.) Married respondents are more likely to hold stock. The propensity to marry and stay married is likely higher for college educated individuals although both could be caused by unobserved personality traits. The impact of children is estimated to be significantly negative which likely reflects parental obligations increasing the importance of liquid savings. We add further controls with the aim of pinning down family influence (conditions in the family when the respondent was in high school) and covariates such as wealth and occupation which are functions of education and cognitive skills. Column (3) includes childhood controls: whether father was a white collar worker, whether the family was well-off, family income in 1957, and father's education. These variables are all significant with the expected signs indicating that children from wealthy, high-income, white collar families are significantly more likely to hold stock later in life—probably because they have learned about financial products from their parents. Because family background is correlated with the decision of going to college, the estimated impact of college declines to 0.106. The likelihood of graduating from college is partly caused by family conditions and partly caused by unobservable factors which affect both family wealth and college going but we cannot narrow this

[30] The results from (unreported) probit regressions are almost identical to the results in Table 8.

[31] To make the results more comparable across different specifications, our regression sample is limited to individuals with non-missing observations on all the relevant controls in our encompassing specification in column (8) of Table 8.

[32] Perhaps, the PSID results on women being less likely to hold stock are due to samples not being representative of all women.

down further.

We next consider cognition in column (4). Cognitive ability is a significant predictor of stockholding, consistent with the findings of Christelis, Jappelli, and Padula (2010) and Cole and Shastry (2009), and including cognitive ability lowers the coefficient to college to 0.10 indicating that cognitive ability after college is a partial channel from college to stockholding. Next, we include occupation. The results in column (5) reveal that white collar workers are significantly more likely to hold stock and, from column (6), this finding is robust to inclusion of a white collar job aspirations of the respondent when in high school. After controlling for occupation the effect of college goes down to 0.082. White collar jobs are more likely to provide retirement accounts such as 401(k) plans which reduce the cost of indirect participation in the stock market which again will lower the informational barriers to direct participation. Our findings are consistent with Hong, Kubik, and Stein (2004): working in a white collar job exposes an individual to a relatively higher number of stockholders which may impact the individual's own choices over stockholding through learning, mimicking, or even the enjoyment of talking about stocks. The results are also consistent with Brown, Ivkovic, Smith, and Weisbenner (2008) who find that individual stockholding is influenced by stock holdings of peers and van Rooij, Lusardi, and Alessie (2011) who document the importance of highly educated peers for individual stockholding. Such peer effects are likely to be correlated with holding a white collar job as colleagues in such jobs are more likely to hold stock and are also likely to be better educated. Which of these more explicit channels—peer effects or learning about the stock market due to indirect participation—dominate the white collar effect, we cannot tell.

Wealth is a highly significant determinant of stockholding and its inclusion lowers the impact of college to 0.071. Wealth is itself a function of stockholding and of college graduation but the inclusion of this endogenous regressor does not change the estimated effect of college on stockholding by much. In the final column, we examine if the type of major matters, specifically if an economics or business degree predicts higher stock ownership as found by Christiansen, Joensen, and Rangvid (2007). The choice of major 40 years or so in the past is clearly not as good a measure of financial literacy as the measures elicited by van Rooij, Lusardi, and Alessie (2011) but it has the advantage of being exogenous to participation in the stock market. We find a positive effect but it is not significant at standard levels of significance.

A strength of the WLS data is the ability to control for family fixed effects. Many parental influences may be unmeasurable or badly approximated by the controls included in the previous table. The WLS interviews siblings of the main respondents (although not for all covariates) and we can therefore control

Table 8: Channels of the Impact of College on Non-retirement Stockholding. Main Respondents. Linear Regressions. Wisconsin Longitudinal data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College dummy	0.159***	0.138***	0.106***	0.100***	0.082***	0.077***	0.071***	0.066***
	(10.14)	(8.13)	(5.98)	(5.57)	(4.34)	(3.86)	(3.63)	(3.26)
Male	–0.046***	–0.043***	–0.040***	–0.039***	–0.045***	–0.044***	–0.054***	–0.057***
	(–3.15)	(–2.92)	(–2.74)	(–2.69)	(–3.05)	(–3.02)	(–3.77)	(–3.91)
Age	–0.034***	–0.029***	–0.029***	–0.029***	–0.028***	–0.028***	–0.028***	–0.028***
	(–3.23)	(–2.72)	(–2.75)	(–2.68)	(–2.62)	(–2.63)	(–2.68)	(–2.66)
Log income 1992	0.007**	0.006**	0.005*	0.005*	0.004	0.004	–0.000	–0.000
	(2.53)	(2.12)	(1.88)	(1.75)	(1.61)	(1.60)	(–0.08)	(–0.08)
Married in 2004	0.133***	0.134***	0.135***	0.136***	0.136***	0.136***	0.112***	0.112***
	(7.37)	(7.39)	(7.52)	(7.55)	(7.58)	(7.58)	(6.25)	(6.24)
Log of number of children	–0.031**	–0.031**	–0.029*	–0.029*	–0.029*	–0.028*	–0.028*	–0.029*
	(–2.02)	(–2.07)	(–1.89)	(–1.94)	(–1.90)	(–1.88)	(–1.91)	(–1.95)
Log of IQ		0.185***	0.138**	0.096	0.074	0.069	0.046	0.043
		(3.33)	(2.46)	(1.60)	(1.23)	(1.13)	(0.77)	(0.71)
Father white collar			0.039**	0.039**	0.037**	0.036**	0.030*	0.029*
			(2.28)	(2.27)	(2.14)	(2.05)	(1.73)	(1.71)
Family well-off			0.044**	0.045**	0.044**	0.044**	0.042**	0.042**
			(2.47)	(2.50)	(2.43)	(2.43)	(2.38)	(2.36)
Log parental inc. 1957			0.036***	0.036***	0.035***	0.034***	0.032***	0.032***
			(3.14)	(3.14)	(2.99)	(2.94)	(2.80)	(2.78)
Father education (years)			0.005*	0.005*	0.004*	0.004	0.005*	0.005*
			(1.81)	(1.77)	(1.69)	(1.64)	(1.78)	(1.80)
Log of cognition in 2004				0.047**	0.041*	0.040*	0.031	0.031
				(2.07)	(1.81)	(1.75)	(1.39)	(1.40)
White collar in 1992					0.050***	0.048***	0.037**	0.036**
					(3.07)	(2.92)	(2.24)	(2.19)
Wanted to be white collar						0.012	0.010	0.011
						(0.68)	(0.59)	(0.62)
Log of wealth in 2004							0.026***	0.026***
							(10.61)	(10.59)
Econ/Business Major								0.034
								(1.25)
Constant	2.626***	1.457*	1.244	1.302*	1.376*	1.410*	1.322*	1.330*
	(3.82)	(1.88)	(1.60)	(1.68)	(1.78)	(1.82)	(1.72)	(1.73)
Adj. R sq.	0.039	0.041	0.050	0.051	0.053	0.053	0.075	0.075
F	35.9	32.6	27.2	25.3	24.0	22.3	31.0	29.2
N	4741	4741	4741	4741	4741	4741	4741	4741

Notes: The left-hand side variable is a dummy equal to 1 if a household owns stock/bonds/mutual funds outside retirement accounts, 0—otherwise. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Robust standard errors.

for family influence by including family fixed effects. The results with family fixed effects are summarized in Table 9. We find that college is still strongly significant but the coefficient is smaller at 0.120—this small drop from the previous table indicates that family background is important but not the main determinant of college graduation and stockholding. The sex of the respondent becomes insignificant, while the effect of IQ is of comparable magnitude but less significant compared to the previous table. IQ matters but, since it also affects cognitive ability, IQ becomes insignificant when the latter is included, while cognitive ability itself is clearly significant. The inclusion of cognitive ability lowers the effect of college to 0.091 consistent with cognitive ability being a channel for college education—having controlled for common family upbringing makes this result more convincing, in our view.

Having a white collar job is highly significant and lowers the coefficient to college to 0.067. White collar job is significant at the five percent level in this table when cognition (which is significant at the one percent level) is included, while white collar dominates cognition in terms of significance in the previous table. Clearly, these variables are correlated as cognitive skills will affect the job one can hold, but the overall picture from the two tables is that both these variables have independent effects. Wealth is still highly significant and lowers the coefficient to college from 0.067 to 0.059 while the inclusion of the economics major dummy drives the education coefficient further down to 0.043, column (6). The dummy for economics and business major is now significant at the one percent level, but as before the impact of the choice of major seems quite independent of the impact of other variables. This variable is not a focus of this study but the result supports the findings of Christiansen, Joensen, and Rangvid (2007) and van Rooij, Lusardi, and Alessie (2011). In column (7), we exclude IQ, cognition, and economics major—the controls we cannot measure in the PSID. As in the IV-regressions using the PSID, college remains significant which points to the importance of college—through improving cognitive abilities—for the decision to hold stocks. (In Appendix B, we run the comparable regression for a PSID sample of siblings, finding similar results although with a larger coefficient to college, likely because of the higher heterogeneity in education in the PSID sample.) Overall, the coefficient to education is to a large extent explained by unobserved family background characteristics. Controlling for this, the results are consistent with college having its main effect through occupational choice, wealth accumulation, and cognitive ability.

Table 9: Channels of the Impact of College on Non-Retirement Stockholding with Family Fixed Effects. Linear Regressions. Wisconsin Longitudinal Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
College dummy	0.120***	0.105***	0.091***	0.067*	0.059*	0.043	0.078**
	(3.72)	(3.14)	(2.69)	(1.89)	(1.67)	(1.20)	(2.25)
Male	–0.003	0.001	0.002	–0.006	–0.013	–0.024	–0.019
	(–0.12)	(0.03)	(0.07)	(–0.22)	(–0.50)	(–0.90)	(–0.71)
Age	0.035	0.043	0.038	0.036	0.031	0.026	0.029
	(1.10)	(1.33)	(1.19)	(1.12)	(0.97)	(0.81)	(0.92)
Age sq./100	–0.025	–0.030	–0.026	–0.025	–0.020	–0.017	–0.020
	(–1.03)	(–1.25)	(–1.09)	(–1.02)	(–0.85)	(–0.69)	(–0.84)
Log income 1992	0.004	0.003	0.003	0.002	–0.001	–0.002	–0.000
	(0.87)	(0.68)	(0.57)	(0.42)	(–0.27)	(–0.38)	(–0.08)
Married in 2004	0.134***	0.136***	0.136***	0.138***	0.125***	0.125***	0.124***
	(4.26)	(4.32)	(4.32)	(4.41)	(3.99)	(4.02)	(3.95)
Log of number of children	–0.058**	–0.061**	–0.063**	–0.059**	–0.061**	–0.060**	–0.057**
	(–2.18)	(–2.30)	(–2.36)	(–2.22)	(–2.30)	(–2.26)	(–2.15)
Log of IQ		0.202*	0.105	0.079	0.057	0.041	
		(1.85)	(0.93)	(0.69)	(0.50)	(0.36)	
Log of cognition in 2004			0.114***	0.109***	0.106***	0.106***	
			(2.92)	(2.79)	(2.73)	(2.74)	
White collar in 1992				0.058**	0.054*	0.055*	0.064**
				(1.98)	(1.84)	(1.90)	(2.20)
Log of wealth in 2004					0.017***	0.017***	0.017***
					(4.41)	(4.41)	(4.46)
Econ/Business Major						0.146***	
						(2.74)	
Constant	–0.739	–1.924	–1.564	–1.375	–1.270	–1.039	–0.737
	(–0.71)	(–1.59)	(–1.29)	(–1.13)	(–1.05)	(–0.86)	(–0.71)
Adj. R sq.	0.023	0.025	0.030	0.032	0.043	0.047	0.038
F	5.7	5.6	5.9	5.8	7.5	7.3	7.6
N	3022	3022	3022	3022	3022	3022	3022

Notes: The left-hand side variable is a dummy equal to 1 if a household owns stock/bonds/mutual funds outside retirement accounts, 0—otherwise. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Robust standard errors.

5 Conclusion

Households with more educated heads are more likely to hold stocks. This is partly due to the more educated having higher wealth but the effect of schooling goes beyond wealth. The level of schooling is partly a function of unobserved ability, attitudes, and taste variables but we isolate the causal effect of college-level schooling by instrumenting it with the number of colleges in the county where and when the household head grew up. We find, using the PSID, a strong positive effect of college education on the propensity to own equity for households with heads who report growing up in families with low or average wealth. Instrumented regressions deliver larger estimates than non-instrumented regressions which indicates a so-called local average treatment effect where college availability disproportionately matters for household heads growing up in less advantaged families combined with college education affecting the tendency to hold stocks more for such heads. The implication is that the size of estimated IV-coefficients cannot be interpreted as the size of the average effect on the full sample.

Using the WLS and non-instrumented regressions, we find that obtaining a college degree clearly affects the probability to own equity and/or bonds and the main channels are occupational choice, wealth accumulation, and cognitive ability.

Appendix A. PSID data

In this appendix, we provide details on construction of consistent measures of the head's race, age, and background variables in the PSID. Race, as well as other demographic information, is recorded at the time an individual enters the survey as head, asked again any other time after an interruption in headship, and asked of every head, new or existing, in 1985, 1990, 1997, and 2005. As a result, race can be recorded multiple times for any head in the sample, and the various records may not agree with each other. We consider the first record of the head's race and ignore records reported by any family member other than the head. (The PSID contains a variable that indicates who is the respondent to the survey questions.) We drop heads who give inconsistent answers about race. We do the same for any other demographic measure that can be recorded multiple times. Background variables that pertain to the head may be recorded differently from one survey to another (for example, after an interrupted headship) if someone other than the head answers the survey questions. Thus, in our selection rules, we "trust" the first record on background variables and any other records provided by the head himself.

The PSID recorded father's education, for household heads, for the first time in 1968 and mother's education in 1974. Individuals report this information the first time they appear as heads of household or after an interruption in the headship. The information was updated for every head, new or existing, in 1997. As a result, heads may have several records on mother's and father's education which may not necessarily agree with each other. To construct a consistent measure of parental education we proceed as follows. First, we keep available records on heads' parental education only if reported by the heads themselves. The parental schooling variables in the PSID are categorical and may be reported with some noise. For example, a head may report that his/her father finished 9–11 grades (some high school) in one year and 12 years (a high school graduate) in some other year. To avoid loss of information due to inconsistent records on parental education, we construct two parental schooling dummy variables (mother and father) equal to one if the head consistently reports that the parent finished at least high school, and zero otherwise.

In 1999 and 2001, the records on parental wealth recall and whether heads lived with both parents while growing up were misrecorded. We therefore ignore records in those years.

The dummy "Rich parents" is set to missing if the records on parental wealth at childhood are inconsistent. We consider the records to be inconsistent if the head claims growing up in a rich family in one survey year and reports growing up in a poor family or family of average well-being some other

year. In a similar manner, we construct consistent measures of whether the head lived with both parents while growing up and whether he grew up on a farm or in a city.

Age in the PSID does not necessarily change in adjacent surveys since information can be collected at different months of a year. Also, some individuals have inconsistent age series which, among other things, may reflect typing errors by interviewers. We utilize information on the year of birth to construct a cleaner measure of age for those heads who have this information in the individual file. Otherwise, we use an individual's age at the time he/she first appears as a head in the survey to impute his/her age in other years.

Appendix B. Channels using PSID data

We ran regressions similar to those reported in Table 9 using PSID data where we control for family fixed effects by matching siblings—see Table B-1. The sample size is smaller than the WLS sample at around 1,500 observations. The coefficient for the college dummy in a regression equivalent to column (1) in Table 9 is 0.13 (0.12 in the WLS) and statistically significant at the one percent level. The coefficients to wealth and the white collar dummy are of magnitudes similar to those found earlier, see columns (2) and (3). However, with PSID data we cannot control for IQ or cognition and the coefficient to college drops to 0.12 when including wealth and occupation controls (relative to 0.08 in column (7) of Table 9 using WLS data). Perhaps, the discrepancy can be explained by measurement error in occupational coding in the PSID as documented in Kambourov and Manovskii (2008). Also, we set the white collar dummy equal to one if the respondent is classified as a manager or administrator (except farm), a sales worker, or a clerical or kindred worker (in the main job) in any of the years observed in the PSID as a head, and zero otherwise. Due to the age variation among PSID respondents, this classification is unlikely to correspond to prime-age as in the WLS. Column (4) in Table B-1 includes an additional finance/insurance/real estate industry dummy in order to proxy for major of study in college, a variable not available in the PSID (real estate workers cannot be separated from finance and insurance professionals in the PSID). Working in those industries increases the probability of owning stock. As in the case of major using WLS data, the effect of this variable seems quite independent of the impact of other variables.

Table B-1: Channels of the Impact of College on Non-Retirement Stockholding with Family Fixed Effects. Linear Regressions. PSID Data

	(1)	(2)	(3)	(4)
College	0.133***	0.124***	0.120***	0.120***
	(3.05)	(2.82)	(2.74)	(2.75)
Male	0.055	0.062*	0.062*	0.066*
	(1.65)	(1.82)	(1.85)	(1.94)
Age	0.014	0.014	0.011	0.011
	(1.21)	(1.23)	(1.00)	(1.02)
Age sq./100	–0.012	–0.012	–0.010	–0.010
	(–0.90)	(–0.90)	(–0.74)	(–0.75)
Married	0.058	0.059*	0.042	0.041
	(1.64)	(1.67)	(1.18)	(1.16)
Log no. of children	–0.028	–0.028	–0.027	–0.028
	(–1.09)	(–1.08)	(–1.04)	(–1.08)
Log income	0.027***	0.026***	0.022***	0.022**
	(3.21)	(3.09)	(2.62)	(2.57)
White collar		0.051	0.052	0.051
		(1.55)	(1.61)	(1.56)
Log wealth			0.007***	0.006***
			(2.93)	(2.89)
Occupation in FIRE				0.049
				(1.24)
Year dummies	Y	Y	Y	Y
Region grew up dummies	Y	Y	Y	Y
F	4.8	4.7	4.9	4.8
N	1530	1530	1530	1530

Notes: The left-hand side variable is a dummy equal to 1 if a household owns stock/bonds/mutual funds outside retirement accounts, 0—otherwise. Regression samples include siblings with non-missing observations on all controls in column (4). "White collar" is a dummy that equals one if the respondent is classified as a manager or administrator (except farm), a sales worker, or a clerical or kindred worker (in the main job) in any of the years observed in the PSID as a head, and zero otherwise. Occupation in FIRE is a dummy for occupation in Finance, Insurance, or Real Estate industries. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. t-statistics in parentheses; robust standard errors.

References

Alan, Sule, 2006, Entry Costs and Stock Market Participation Over the Life Cycle, *Review of Economic Dynamics* 9, 588–611.

Ameriks, John, Andrew Caplin, and John Leahy, 2003, Wealth Accumulation and the Propensity to Plan, *Quarterly Journal of Economics* 118, 1007–1047.

Andersen, Steffen, and Kasper Meisner Nielsen, 2010, Participation Constraints in the Stock Market: Evidence from Unexpected Inheritance due to Sudden Death, *Review of Financial Studies, forthcoming.*

Attanasio, Orazio P., James Banks, and Sarah Tanner, 2002, Asset Holding and Consumption Volatility, *Journal of Political Economy* 110, 771–792.

Brown, Jeffrey R., Zoran Ivkovic, Paul A. Smith, and Scott Weisbenner, 2008, Neighbors Matter: Causal Community Effects and Stock Market Participation, *Journal of Finance* 63, 1509–1531.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, 2007, Down or Out: Assessing the Welfare Costs of Household Investment Mistakes, *Journal of Political Economy* 115, 707–747.

Cameron, Stephen V., and Christopher Taber, 2004, Estimation of Educational Borrowing Constraints Using Returns to Schooling, *Journal of Political Economy* 112, 132–182.

Campbell, John, 2006, Household Finance, *Journal of Finance* 61, 1553–1604.

Card, David, 2001, Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica* 69, 1127–1160.

Christelis, Dimitris, Tullio Jappelli, and Mario Padula, 2010, Cognitive Abilities and Portfolio Choice, *European Economic Review* 54, 18–38.

Christiansen, Charlotte, Juanna Schroter Joensen, and Jesper Rangvid, 2007, Are Economists More Likely to Hold Stocks?, *Review of Finance* 12, 465–496.

Cole, Shawn, and Gauri Kartini Shastry, 2009, Smart Money: The Effect of Education, Cognitive Ability, and Financial Literacy on Financial Market Participation, *Working Paper, Harvard Business School* 09-071.

Currie, Janet, and Enrico Moretti, 2003, Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings, *Quarterly Journal of Economics* 118, 1553–1604.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, 2010, Are Risk Aversion and Impatience Related to Cognitive Ability?, *American Economic Review* 100, 1238–1260.

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

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaai, 2011, IQ and Stock Market Participation, *Journal of Finance* 66, 2119–2164.

Guiso, Luigi, Tullio Jappelli, and Daniele Terlizzese, 2005, Income Risk, Borrowing Constraints, and Portfolio Choice, *American Economic Review* 86, 158–172.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2008, Trusting the Stock Market, *Journal of Finance* 63, 2557–2600.

Haines, Michael R., 2004, Historical, Demographic, Economic, and Social Data: The United States, 1790-2000 [computer file], Inter-university Consortium for Political and Social Research, ICPSR02896-v2.

Haliassos, Michael, and Carol C. Bertaut, 1995, Why Do So Few Hold Stocks?, *Economic Journal* 105, 1110–1129.

Hansen, Karsten T., James J. Heckman, and Kathleen J. Mullen, 2006, The Effect of Schooling and Ability on Achievement Test Scores, *Journal of Econometrics* 121, 39–98.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social Interaction and Stock-Market Participation, *Journal of Finance* 59, 137–163.

Hryshko, Dmytro, María José Luengo-Prado, and Bent E. Sørensen, 2011, Childhood Determinants of Risk Aversion: The Long Shadow of Compulsory Education, *Quantitative Economics* 2, 37–72.

Kambourov, Gueorgui, and Iourii Manovskii, 2008, Rising Occupational and Industry Mobility in the United States: 1968-97, *International Economic Review* 49, 41–79.

Lovenheim, Michael F., 2011, The Effect of Liquid Housing Wealth on College Enrollment, *Journal of Labor Economics* 29, 741–771.

Perraudin, William R.M., and Bent E. Sørensen, 2000, The Demand for Risky Assets: Sample Selection and Household Portfolios, *Journal of Econometrics* 97, 117–144.

Stock, James H., Jonathan H. Wright, and Motihiro Yogo, 2002, A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics* 20, 518–529.

van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie, 2011, Financial Literacy and Stock Market Participation, *Journal of Financial Economics* 101, 449–472.

Vissing-Jørgensen, Annette, 2002a, Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy* 110, 825–853.

——— , 2002b, Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures, NBER Working Paper # 8884.

Yao, Rui, and Harold H. Zhang, 2005, Optimal Consumption and Portfolio Choices with Risky Housing and Borrowing Constraints, *Review of Financial Studies* 18, 197–239.