

NBER WORKING PAPER SERIES

RACIAL DISPARITIES IN SAVINGS BEHAVIOR FOR A CONTINUOUSLY EMPLOYED
COHORT

Kai Yuan Kuan
Mark R. Cullen
Sepideh Modrek

Working Paper 20937
<http://www.nber.org/papers/w20937>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2015

We would like to thank Liran Einav for his helpful comments and Alex Christodoulou for helping us assemble the data for this project. This work was supported by several grants including a supplement from the National Institute On Aging of the National Institutes of Health under Award Number R01AG026291 and the Stanford Center for Longevity. The findings and conclusions are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health, the Stanford Center for Longevity, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Kai Yuan Kuan, Mark R. Cullen, and Sepideh Modrek. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Racial Disparities in Savings Behavior for a Continuously Employed Cohort
Kai Yuan Kuan, Mark R. Cullen, and Sepideh Modrek
NBER Working Paper No. 20937
February 2015
JEL No. D14,D31,I31,J11,J32

ABSTRACT

The wealth gap has reached record highs. At the same time there has been substantial proliferation of 401(k) savings accounts as the dominant retirement savings vehicle, and these accounts make up an increasing proportion of overall wealth. In this paper we examine 401(k) saving behavior of continuously employed workers over an eight-year period at a single, geographically diverse employer. We demonstrate substantial difference in 401(k) savings behavior by employee ethnicity even within a single employer 401(k) plan architecture. We show both African American and Hispanic employees are less likely to participate in the 401(k) plans. Moreover, conditional on participation African Americans contribute a lower proportion of their income to their 401(k) plan on average. We also show that African Americans and Hispanics tend to draw down on their 401(k) balances more often. Finally, we document that both African Americans and Hispanics favor safer assets within their plan options. Together these differences substantially impact the level of 401(k) balances accumulated and therefore overall wealth accumulation.

Kai Yuan Kuan
Stanford University
Statistics Department
390 Serra Mall
Stanford, CA 94305-4065
kykuan@stanford.edu

Sepideh Modrek
Department of General Medical Disciplines
School of Medicine, Stanford University
1070 Arastradero Road Room 309
Palo Alto, CA 94304-1334
smodrek@gmail.com

Mark R. Cullen
Stanford University School of Medicine
1265 Welch Rd X338
Stanford, CA 94305
and NBER
mrcullen@stanford.edu

Introduction

The wealth gap between whites and minorities in the United States has reached record highs (Taylor, Kochhar et al. 2011, Shapiro, Meschede et al. 2013). Several studies document the importance of higher incomes, larger inheritance and higher rates of home ownership amongst whites in the widening gap in the wealth accumulation (Gittleman and Wolff 2004, Borgo 2013, Shapiro, Meschede et al. 2013), but controversy remains over the role of savings behavior and asset returns in accelerating the gap.

Savings behavior, the choice to save or spend a certain proportion of one's income and how to save across investment options (i.e. housing, pensions, stocks, bonds, etc), is very difficult to assess because detailed longitudinal data is required. Though several studies have tried to assess racial difference in asset accumulations and returns, some controversies remain. Gittleman and Wolff (2004) use the Panel Study of Income Dynamics (PSID) data, which follows the same families to examine differences in the rate of wealth accumulation of white Americans relative to African Americans. They find that the majority of the difference in asset accumulations over the 10-year period is driven by larger inheritances and incomes for white families. They find no difference in savings rate conditional on income and find no evidence of difference in asset returns. In a more recent study Del Bago (2013) uses the Health and Retirement Survey and finds that after accounting for income differences some of the wealth gap can be explained by lower capital gains for African Americans. She finds that the lower capital gains may be due to the lower appreciation rate of housing owned by African Americans and differences in the rate of investments in high-return assets such as stocks. Another study by the Social Security Administration (Choudhury 2001), which uses administrative data, finds that at every income quartile and educational level, the percentage of African American and Hispanic households that own risky, higher-yielding assets is considerably smaller than the percentage of white households that do. She concludes that some of the wealth gap appears to be attributable to differences in savings behavior and rates of returns on investments chosen by each ethnic group.

Another major barrier to understanding the wealth gap is that incentives to save in non-housing wealth, particularly 401(k) type pensions, vary substantially by one's occupation and even one's company. Therefore, even if studies can control for income, they may not be able to adequately account for the vastly different savings incentives offered by employers. As employer-based defined contribution plans replace traditional defined benefits pensions, these saving vehicles are becoming a larger share of wealth; understanding racial and ethnic differences in these savings vehicles becomes increasingly important because individuals bear the responsibility of ensuring adequate saving for their retirement.

This study examines one aspect of asset accumulation and savings behaviors, using company-sponsored 401(k) benefits. Following a cohort of workers from a single, geographically diverse national company will allow us to examine the extent to which savings behavior differences persist in participation, contributions, balances, and asset allocations in 401(k) pensions across different minority groups facing similar information and incentives. Importantly we intend to control for changes in health (Pollack, Cubbin et al. 2013), document differences in choices of asset types across groups and comment on pathways, which may underlie the observed differences and ultimately accelerate the wealth gap.

Factors that May Affect 401(k) Behavior

Here we survey the relevant literature in preparation to attempt to clarify potential pathways via which any observed differences in behavior across ethnicities may manifest.

401(k) Plan Awareness and Financial Literacy

An individual's familiarity with the structure of the 401(k) plan offered by his/her employer and understanding of the various investment options may be an important factor in determining his/her willingness to participate

and contribute to 401(k) plans. Familiarity breeds confidence, and individuals who feel that they understand how the system works may be more willing to participate and also trust a larger portion of their savings to the plan. More financially-experienced individuals are also likely to allocate funds within a pension plan differently.

Clark and Schieber (1998) examine the effect of plan awareness on 401(k) participation and contributions via a dataset of employees from 17 different companies ranging in size from 7,000-10,000 workers. They separate the workers via inferred different levels of employee communication. At the first level, workers received only plan enrollment forms and periodic statements of balances. At the second level, workers received generic newsletters regarding information on 401(k) plans. The third level of communication involves communication materials specifically tailored to the employer's 401(k) plan. They find that more dissemination of information increases participation and contributions. In a more recent study, Shah Goda, Manchester et al. (2012) uses an experimental design to provide over 16,000 employees at the University of Michigan customized information on projected future account balances and retirement income bases on increases in current participation and contributions. Using administrative data, they find increased participation and contributions for those who receive the most informative and customized mailings.

Beyond specific plan knowledge, studies have also shown that general financial literacy (i.e. such as a working understanding of the benefits of compound interest, risk diversification, tax treatment of retirement savings, and inflation) has some effect on 401(k) behavior. Mitchell and Hastings (2011) examines the linkage between financial literacy and retirement savings using evidence from the 2009 Chilean Encuesta de Proteccion Social (EPS) (similar to the US Health and Retirement Survey). They find that financial literacy has a significant and positive effect on pension savings controlling for education and income. This effect appears to hold in the US as well. Using US Health and Retirement Survey data, Lusardi and Mitchell (2007) find a positive link between financial literacy and planning, as well as between planning and wealth. They show that planning increases wealth rather than vice-versa and again conclude that financial literacy has a positive effect on retirement savings. Cole and Shastry (2009) find a positive relationship between financial literacy and general financial market participation (Heinberg, Hung et al. 2014). However, while participation is increasing in financial literacy, they find no significant effect on portfolio allocations.

Important for this study, several previous studies document substantial differences in financial literacy amongst minorities as compared to whites. Indeed, African Americans and Hispanics display the lowest level of financial knowledge in the U.S. context, regardless of the financial literacy measure used (Lusardi and Mitchell 2014).

Institutional (Dis)Trust

A complementary literature finds important differences in attitudes regarding saving institutions between white Americans and minorities (Lusardi and Mitchell 2014). Studies suggest that African Americans may generally feel mistrust in investing in equity markets (Mabry 1999), or perceive discrimination and self-select away from formal financial institutions where they feel discriminated against (Longhofer 2004, Apgar and Calder 2005). Recent media accounts highlight historical and recent accusation of racial discrimination on the part of financial institutions (Savage 2011, Savage 2012, Coates 2014). These accounts suggest that distrust is justified and may further exacerbate latent distrust. Alternatively, difference may be driven by the access to information and the types of financial institutions that serve different ethnic groups which are often segregated in different communities (Rugh and Massey 2010).

Network Effects

Many aspects of human behavior are now believed to be influenced by peer groups (Evans, Oates et al. 1992, Karlan 2007, Webel, Okonsky et al. 2010), and there are also studies suggesting that 401(k) behavior is as well. Duflo and Saez (2002) examine the pension decisions of employees of a large university. They identify peer groups by dividing individuals into sub-groups by departments, and instrumenting average group participation using salary and tenure within groups, find a significant effect of peer behavior on participation and fund

choice. Outside the realm of pensions, Hong, Kubik et al. (2005) find that stock market participation increases with peer participation. If minorities have historically been less likely to invest in financial assets, then peer effects could further explain current differences.

Investor Economic Preferences

The investment goals of the individuals in question may lead to differing 401(k) contribution behavior. Munell, Sunden et al. (2001) examines the effects of an individual's planning horizons on contributions and participation using data from the Survey of Consumer Finances (SCF). They find that individuals with longer planning horizons tend to participate and contribute more. Relatedly, Mitchell and Hastings (2011) examine the effect of (im)patience on contributions using evidence from a dataset of Chilean EPS respondents (see financial literacy section). They find that individuals who are better able to forgo current gratification for a better future outcome tend to also place more of their income into pensions accounts.

Another important investor preferences that may relate to 401(K) savings is risk aversion. While most 401(k) plans allow participants to invest in low risk options, participants may not be fully aware of this or may not feel comfortable taking on the risk associated with a somewhat illiquid asset. The theoretical literature suggests that risk preferences influence stock market participation, contributions and allocations (Michaelides and Gomes 2005), and seminal work from the Health and Retirement Survey on risk aversion provides empirical evidence that risk aversion affects a series of risk related behaviors including investing in stock (Barsky, Kimball et al. 1997).

Non-401(k) Financial Circumstances

Financial theory suggests that 401(k) savings decisions and asset allocations within 401(k) accounts are not made in a vacuum. Rather, investors consider all their asset holdings when making financial plans. Many issues can influence an individual's pension savings, including spousal income and savings behavior, existing wealth, medical conditions and family expenditures.

Pence (2002) suggests that 401(k) savings have little to no effect on actual household savings. Rather, she finds weak evidence that individuals substitute external (non-401(k)) real assets for 401(k) balances. If individuals have marginal utility functions decreasing in wealth, we would expect that individuals with higher wealth would save less. If utility is also decreasing in absolute risk aversion (DARA), we would expect individuals with higher wealth to maintain riskier (in the absolute sense) portfolios. Given that minorities have lower wealth holdings than do whites, and also have lower incomes (Gittleman and Wolff 2004), this would suggest that all other factors equal, a minority individual may behave differently to a difference in wealth (either inherited or from previous income).

Household Budget Constraint

While current individual income will obviously affect participation, contributions and balances in 401(k) accounts, household income is a more important determinant. Households jointly optimize their retirement savings accounting for relative current resources and constraints as well as desired long-term savings (Lundberg and Ward-Batts 2000). Current household income, which would include any spousal income, and current total retirement savings, which would include spousal saving, are key factor in determining a worker's individual choices in their personal 401(k) account. In addition family composition affects individual savings through their budget constraint and could affect substantially saving behavior (Scholz and Seshadri 2007).

There are well-documented differences in household income, particularly for women across racial and ethnic lines (Corbetti and Hill 2014), that could drive any relationship between race and savings. There are also differences in household composition, which should affect available resources (Passel, Livingston et al. 2012).

Plan Characteristics

Employer 401(k) plans differ along two key dimensions: employer match rates (the percentage of employee contributions that the employer contributes on the employee's behalf) and access to funds (the ability draw upon 401(k) balances before retirement via withdrawals and loans (Love 2007).

In theory the existence and introduction of a match rate has an ambiguous effect. On one hand, a larger match rate increases the return on 401(k) savings, but at the same time allows individuals to meet retirement targets with lower contributions. The effect of employer match rates on contributions depends on which of these two effects dominates. Evidence suggests that access to funds and the introduction of an employer match increase participation and contribution (Basett, Fleming et al. 1998, Clark and Schieber 1998, Munell, Sunden et al. 2001). However, the effect of introducing an employer-matching contribution is small on non-highly compensated workers (Mitchell, Utkus et al. 2005). The effect of increasing an existing employer match rate is controversial and previous studies have encountered mixed results. While some researchers have found that increasing the match rate increases contributions (Pepke and Poterba 1995, VanDerhei and Copeland 2001, Choi, Laibson et al. 2002, Engelhardt and Kumar 2007, Chetty, Friedman et al. 2012), others have found the opposite (EBRI 1994, Munell, Sunden et al. 2001).

The option to access funds in 401(k) savings plans should theoretically affect participation rates and contributions by making such assets more liquid in case of an adverse economic shock. The IRS sets a 10% penalty in addition to the income tax ramifications for early withdraws from 401(k) savings accounts. However, the IRS mandated penalty does not apply in many cases, and many early withdrawal events are not penalized. The most important exception is for employees aged 55 or older who take a withdrawal when they separate from an employer. Taxpayers can also get an exemption from the penalty if they take withdrawals for hardship reasons, home purchase, and post-secondary education (Argento, Bryant et al. 2015). Employee access to 401(k) accounts via loans is regulated by the employer's particular plan instead of the IRS. VanDerhei, Holden et al. (2009) calculate that 88% of 401(k) participants belong to plans that allow participant to take loans, and the terms of the loan are set by individual savings plans. These loan characteristics have previously been shown to affect the amount and frequency of loans (Beshears, Choi et al. 2010).

In Figure 1 we summarize these overarching pathways. We denote pathways for which we believe race differences could be an important contributor. We also highlight those pathways for which we have some measure that we could exploit to better understand pathways to explain the racial difference we document.

Methods

We examine 401(k) savings and withdrawals behavior of employees over a period of 8 years (most of our models only include 6 years due to missing data for 2008 and data issues in 2005), specifically focusing on differentials between Whites, African Americans and Hispanics).

We examine savings behavior along five dimensions for each year in the sample period. These include participation in the employer's 401(k) plan, and conditional on participation, we examine 401(k) contributions, balances, withdrawals/loans and whether an individual made any contributions to a particular kind of fund in that year (stock funds, money markets and bond funds). During each year, we observe information about the individual's salary, health conditions as well as detailed demographics about each person.

This setting allows us several advantages in addressing the racial wealth gap. First, we are able to examine individuals who are making decisions within a common incentive system. This allows us to overcome the issue of heterogeneity in employer-offered 401(k) plans, which may otherwise interact with racial distribution across employers in a way that could bias the analysis. Second, the heterogeneity among individuals in our dataset is likely reduced compared with most other analyses in the literature, which sample nationally and across a wide array of occupations. While reducing the generalizability of the results obtained, this also allows us to reduce the error resulting from widely differing demographic characteristics of different races across occupations.

Finally, we are able to observe the behavior of our cohort with relatively little measurement error because it is derived from administrative records, as opposed to self-reported savings, which has been previously shown to be inconsistent (Gustman, Steinmeier et al. 2008). These data are derived from the company's third-party 401(k) administrator and therefore provide us with precise numerical data for contributions, balances and most of the demographic variables. In addition, we are able to conduct a variety of mediation analyses to understand pathways underlying the observed differences. Despite these strengths, there are still many factors that we are unable to fully account for in this analysis, e.g. non-401(k) financial circumstances, institutional trust, network effects described above. Most notably we have no information on personal or family savings external to the employer's 401(k) plan and household income and assets. Furthermore, given the complexity surrounding the factors that drive savings behavior, which include cultural, historical and demographic factors, we focus on describing the differences and test several hypotheses regarding the underlying pathways.

Data

Baseline Sample

In order to ensure that we observe each individual's behavior for all the years in our sample period, we restrict our sample to those who are continuously employed over the period 2003-2010. The sample of workers survived a series of large layoffs in 2009, and therefore previous studies have shown that this surviving sample is selected to be healthier (Modrek, Hamad et al. 2015) and likely more productive. This does not represent a major problem as the main focus of this paper is behavior differentials between ethnicities; reducing the sample in this manner only serves to illustrate that the results we obtain hold across individuals of possibly higher levels of financial literacy, albeit with reduced generalizability.

We make three further restrictions on our cohort:

- i) Employee type: We remove executives from the sample as they are not likely to behave the same way as regular employees. (205 individuals removed from the panel.)
- ii) Salary: We retain only workers who earned at most \$200,000 for every year in the sample period. This avoids bias resulting from the cap in 401(k) contribution amounts (see plan details below) and differences in behavior of especially high earners, who are also likely to be more financially literate. This represents 1.1% of the sample, or 115 people. Furthermore, individuals who have high salaries are more likely to participate in outside investment vehicles and otherwise have more savings outside the plan (Bricker, Kennickell et al. 2012). Since we are unable to account for external investments, restricting our sample to those who do not earn too much reduces the bias from such considerations.
- iii) Risk Score: This is a numerical score (produced by Verisk®) designed to capture the amount an individual is expected to pay for health care in a given year. It is normalized such that a risk score of 1 corresponds to a mean level of health care payments (with respect to the population of US workers covered by commercial health insurance in that year). A risk score of x in a year then indicates that the individual is expected to pay x times the mean for health care in the following year. In order to reduce bias from extremely ill individuals, we limit our cohort to only individuals whose risk score never exceeds 6 during the 2003-2010 period. (461 people removed, not including the 115 individuals removed as a result of the salary limit)

Our final baseline panel comprises 9,622 workers and 71,386 observations over 7 years, and will be further reduced depending on availability of data (for most of our analyses we did not receive data for 2008, and we have excluded 2005 from the analysis as there are some data irregularities in that year that we have been unable to account for). Table 1 provides demographics for this cohort. The cohort is predominately male, 83%, and about three-quarters hourly workers across 29 different US locations. The racial composition of the cohort is 83% White, 8.3% African American, 6% Hispanic and 2.8% Other (this category includes Native Americans and Asian Americans). The average age in our final observation year was 52. The median annual salary was \$53,666 over all years in the sample period and the median risk score was 0.788 (our sample is healthier on

average than the population of working US adults covered by employer sponsored health insurance). Our cohort held a median of \$52,522 in 401(k) balances.

Outcome Measures

Participation

Our participation measure is a binary variable, which takes on the value 1 if workers participated in the 401(k) plan. If any contribution was made to the 401(k) account, either by the individual or the company in a given year, then the worker is considered to be a participant. In Figure 2 we present the secular trend in participation rates by age and race (including only the three largest race/ethnic groups) over the study period. For each group there is a secular trend of increasing participation over the eight-year period, but with some decline or stabilization after 2009.

Contributions

We measure contributions in two ways, either as the log of contributions or as a percent of the worker's yearly income. The former is the natural log of the dollar amount contributed by the individual to all the 401(k) accounts in the given year. Transforming monetary values by taking logs is customary when dealing with dependent variables that are highly skewed to stabilize the variance. The latter is, as its name suggests, the proportion of an individual's yearly income that was made in contributions to his/her 401(k) accounts. We include in the regressions only those individuals who participated, and we use White's heteroskedasticity correction as an added measure against inaccurate standard errors. In both measures the contributions variables include both contributions made by the employee herself and those made by the employer on her behalf.¹ Figure 3 displays median contributions by age and race, showing an increase in median contributions over time, as well as a tendency for White workers to contribute more than their Hispanic or African American counterparts of the same age group.

Balances

We measure total balances in all the 401(k) accounts on one date each year. Here again we log transform the balances to stabilize the variance of the dependent variable. Figure 4 illustrates the median balances held by individuals over time, by age and race. As expected, the balances grow over time, with some widening of the gap between groups.

Withdrawals and Loans

Hardships or large purchases may cause an individual to require access to a substantial sum of money during some point in his or her career. The employer's 401(k) savings plan offers withdrawals and loans as methods for employees to gain access to funding sourced from their 401(k) savings and they differ only in the technicalities of when an individual may request one or the other as well as the penalties that apply. As such, we analyze them together and to provide a joint picture of how often individuals of different races access their 401(k) savings during their working lives. There were a total of 10,151 withdrawals and 7,915 new loans made during the sample period.

Individuals 59.5 years or younger may at any time withdraw the current market value of their after-tax contributions, whereas pre-tax contributions can only be withdrawn under specific conditions such as serious medical conditions, to pay tuition for graduate or college education for oneself, spouse or legal dependents, or for funeral expenses, and only after other avenues of sourcing funds are exhausted, including taking a loan on 401(k) balances. There are no restrictions on withdrawals for individuals older than 59.5. There may be tax

¹ While we are unable to separate out the contributions made by the employer and those made by the employee for most years, we are able to do so for a smaller set of years and this is included as a robustness check detailed below.

penalties associated with early withdrawal and both income and capital gains taxes are due on untaxed income that is withdrawn.

Loans from the 401(k) plan are divided into general purpose loans and primary residence loans. An individual with a 401(k) savings account may request a general purpose loan at any time for any purpose, whereas as one would guess, a primary residence loan is available only to purchase a home. Loans act as vehicles for an individual to withdraw their balances as they are treated as loans taken from the individual's 401(k) account itself; an individual cannot borrow more than 50% of their remaining 401(k) account balances (minus any outstanding loans), interest is charged at the prime interest rate and any repayments and interest is accrued to the individual's own 401(k) account. Any defaults are treated as withdrawals.

Figures 5 and 6 graph the rate at which new withdrawals and loans are taken out by race and age. Within a given age group, African Americans tend to make the most withdrawals and in most years take the most loans as well, followed by Hispanics. Whites persistently tend to take out the least loans and make the least withdrawals in a given year.

Allocations

Several studies have shown that one of the potential drivers of the ethnic wealth gap is the way in which people of different ethnicities invest their money. In this set of models we examine how people of different races choose to allocate their savings between various asset classes. Individuals within the employer's 401(k) plan are offered a range of choices in terms of where to invest their 401(k) savings. Choices comprise funds run by various providers, and differ in terms of their risk/return objectives and maturities. In order to abstract from these complexities, we separate the menu of choices into three categories: stock funds, bond funds and money market funds. Employees are also offered company stock as an investment category, but we choose to exclude all investments in this category as during the sample period, the 401(k) plan was such that all employer match contributions were made to company stock, and a holding period was required before funds held in company stock could be rebalanced into other asset classes (this policy has since been changed). This makes it difficult to measure, which contributions into company stock are intentional and which are simply due to company policy.

The set of models used here are logistic regressions, where the dependent variables are binary and take on the value 1 if an individual makes a contribution to the given type of fund in a particular year. By using binary as opposed to continuous contribution amount variables, we avoid details regarding percentage portfolio allocations and differing levels of 401(k) savings and simply focus on whether or not an individual made an investment in a particular asset type. Unless the individual has substantial outside asset holdings, we expect individuals to diversify their holdings and invest in various asset classes.²

Statistical Methods

Primary Analyses

Our primary analyses examines saving behavior along four different dimensions: participation, contributions, balances, and withdrawals/loans. For each outcome, we estimate a series of regression to describe differences by race/ethnicity. In our base model, model 1, we estimate

² We assume that middle to upper-middle income earners in our sample will likely hold a large proportion of their financial market investments in their 401(k) accounts as opposed to outside vehicles (such as private brokerage accounts) due to the favorable tax treatment received by 401(k) investments as well as the generous terms of the company's 401(k) plan. Reasoning along these lines we would expect that an analysis relying on binary variables would not have much to say about the issue of racial investment allocations but surprisingly, we will show strong pattern in the classes of assets minorities and Whites invest their 401(k) savings in.

$$Y_{lit} = f(\beta_1 R_i + \beta_2 X_i + \beta_3 I_{lit} + \beta_4 H_{lit} + \beta_5 U_{lt} + \beta_6 L_l + \beta_7 T_t + \epsilon_{lit})$$

where the Y_{lit} is the savings outcome of interest, R_i is an indicator variable for race, X_i is vector of time-invariant individual factors such as birth year, gender and employee type, I_{lit} is yearly income from company W2 filings, H_{lit} is the individual's risk score in a specific year, U_{lt} is the county-level yearly unemployment rate, L_l is series of controls for each location, and T_t is a series of controls for year. For dichotomous outcomes these regressions are modeled using logistic random effects models and for continuous outcomes they are modeled using random effects linear regressions. All models account for heteroskedasticity by using Huber-White (White 1980) robust standard errors. We similarly examine differences 401(k) portfolio allocation mix by race. In this set of models we examine how people of different races choose to allocate their savings between various asset classes categorized into classes as noted above. For each of these asset classes, we estimate three models, similar to model 1, where Y_{lit} is dichotomous variable to indicate if an individual invested any portion of their portfolio in that specific asset class.

Sensitivity Analyses

We also conduct two supplemental analyses to ensure that the measurement of the participation or contribution outcome is not sensitive to the method of measurement. For participation, we note that some may consider it more appropriate to limit participation to individuals who actively participate, i.e. if the individual himself/herself made any contribution to the 401(k) account. In several years of the data we cannot separate employee or employer contributions; however, we can do so for five of the survey years. Therefore we re-estimate Model 1 defining participation as only those who make a contribution themselves.

We also re-examine the measurement of contributions. One potential source of error in this analysis is the structure of the 401(k) plan, which allows contributions both before and after tax. Before-tax and after-tax contributions are both recorded as their respective dollar amounts in our dataset, but are treated differently for accounting purposes in order to avoid double taxation. Aside from that, all 401(k) plans also stipulate different limits to the maximum allowable contribution depending whether contributions are made before or after tax. Therefore we reexamine contributions only made before taxes since so few people have after tax contributions.

Pathway analyses

In order to better understand our descriptive results, we perform mediation analyses by examining the same models as above, but adding one more construct, the mediator, to see if we can attenuate the relationship between race and saving behavior.

(Model 2)

$$Y_{lit} = f(\beta_1 R_i + \beta_2 X_i + \beta_3 I_{lit} + \beta_4 H_{lit} + \beta_5 U_{lt} + \beta_6 L_l + \beta_7 T_t + \beta_8 M + \epsilon_{lit})$$

This model is identical to model 1, except for the inclusion of, M , the mediating variable. We are interested if its inclusion reduces the magnitude of the race relationship estimated by β_1 . We investigate a series of mediators including family composition (Scholz and Seshadri 2007), risk aversion, financial literacy (Lusardi and Mitchell 2007) and network effects (Duflo and Saez 2002). Due to missing data, the mediation analysis will use a subset of observation and the number of observations will vary in each of these analyses. Again, for dichotomous outcomes these regressions are modeled as a logistic random effects model and for continuous outcomes they are modeled using random effects linear regressions. All models account for heteroskedasticity by using Huber-White robust standard errors.

Results

In this section we will focus on describing the results from our main set of analyses and on addressing the various forms of errors that could arise as a result of variable definitions and the setting in which the data were generated and recorded. In the next section, we will discuss and test various hypotheses that could underlie the results presented in this section.

Participation, Contributions and Balances

Table 2 presents results from our base model on participation, contributions, and balances. Notably, even among this fairly homogeneous sample of employees there is a significant and substantial reduction in the likelihood that a given African American or Hispanic individual would participate in the plan. African American are least likely to participate in employer sponsored 401(k) plans (OR = 0.26, 95% CI = [0.18, 0.39]) compared to White employees of similar background. Similarly Hispanic employees are also less likely to participate in the 401(k) plan than their White colleagues, (OR = 0.36, 95% CI = [0.23, 0.57])³ Columns 2 and 3 of Table 2 presents the outcome of linear regression analyses where the dependent variables are logged contributions and the 401(k) contribution rate as a percent of annual income respectively. The analysis in Table 2 column 3 suggests that conditional on participation, Hispanics contributed a similar proportion of their income their 401(k) savings accounts as Whites. African Americans on the other hand, participate less and moreover, those who do participate tend to contribute less. The coefficients suggest that African Americans who do invest contribute 1.9% less of their income than do Whites ($\beta = -1.86$, 95% CI = [-2.25, -1.48]). Column 4 presents the results of a linear regression on logged 401(k) balances, limited to individuals who participate in the plan. The coefficients in this analysis indicate that both African Americans and Hispanics maintain lower 401(k) balances overall than Whites.

Withdrawals and Loans

Table 3 presents the regression analyses of race and a set of explanatory variables on withdrawals and loans made on 401(k) balances. For each of these binary dependent variables, we use a logistic regression to model the outcome. We limit the regression to only individuals who participated in the 401(k) plan. African Americans within in our sample are much more likely to make a withdrawal (OR = 8.2, 95% CI = [6.2, 10.9]) and to take a loan (OR = 2.2, 95% CI = [1.9, 2.5]) in a given year than White employees.⁴ Hispanics also withdraw at a higher rate than do Whites (OR = 1.83, 95% CI = [1.3, 2.6]), and also take out loans at a higher rate (OR = 1.7, 95% CI = [1.4, 2.0]). Both Hispanics and African Americans draw down their 401(k) balances through withdrawals and loans at a higher rate than Whites, but African Americans take withdrawals at a much higher rate than even their Hispanic colleagues. Drawing down on 401(k) balances not only reduces retirement savings by the amount withdrawn (although loans are repaid and so the net effect is that the principal remains in

³ As a benchmark, the odds of a randomly selected individual participating in the plan was 4.75 during 2003, the year with the least participation and 11.30 during 2007, the year when participation was the highest. In order to get a sense of the magnitude of this, if the odds ratio were at those levels, then a ballpark estimation indicates that all else being equal the probability of an African American participating would be reduced to 53% from 83% in the former case, and 73% from 92% in the latter.

⁴ As a ballpark estimate to illustrate the impact of an African American coefficient of 8, using as a baseline probability the number of withdrawals (11,663) as a proportion of total observations in the sample (67,354), we get a baseline probability of 17%. This gives an odds ratio of 0.21, and the African American flag increases this odds ratio to 1.68, thus corresponding to a probability of 63%. Given that we have 1746 observations of African Americans making withdrawals in our sample, and 4452 person-year observations of African Americans in total, this translates to 40% probability of a randomly selected African American making a withdrawal in a randomly selected year. This compares to a 17% probability of a randomly selected White person making a withdrawal in a randomly selected year. The case for loans is similar, and the effect is small only in comparison to the effect on withdrawals.

retirement savings), but also causes the individual to miss out on any capital gains that might have resulted had those balances been held within the 401(k) account.

Allocation-Mix Models

In Table 4, we present difference in allocation mix by race. We find that relative to Whites, Hispanics and African Americans in our sample tend to favor safer asset classes as opposed to stocks. We see that both Hispanics and African Americans contribute to stock funds at a lower rate than their White colleagues. African Americans are about half as likely as Whites to contribute to stock funds (OR = 0.43, 95% CI = [0.26, 0.70]), (OR = 0.43, 95% CI = [0.26, 0.70]), as are Hispanics (OR = 0.49, 95% CI = [0.27, 0.90]). These minorities seem to contribute to stocks similarly. At the same time, both African Americans and Hispanics are more likely to contribute to money market funds than their White colleagues. African Americans contribute to money market funds with about twice the odds of Whites (OR = 2.85, 95% CI = [1.92, 4.22]). Likewise, Hispanics also contribute to money market funds similarly to African Americans but much more likely than Whites (OR = 1.95, 95% CI = [1.21, 3.15]).

Robustness Analyses

Alternative definitions of participation

Some may argue that it is more appropriate to consider the individual an active participant only if the individual herself made the contribution. While there were data-related challenges in separating out individual contributions from employer contributions for all years, for 5 years (2003, 2004, 2008, 2009 & 2010) we were able to separate out employee from employer contribution. We re-estimate Model 1 using only self-participation as the outcome variable. We find almost identical similar magnitudes for the lower participation rates of African Americans and Hispanic employees relative to White employees (See Table 5, columns 1 and 2). Therefore, we conclude that this definition distinction does not affect our results.

Pretax contributions

As noted above there is different marginal benefits to 401(k) contributions based on whether the contribution is before tax or after. In the analyses above we present all contributions. However, only about 10% of the cohort makes any after tax contributions to the 401(k) plans. To ensure the robustness of our results, we further verify that our results are indeed robust to examining only pre-tax contribution, by including a separate regression with only before tax contributions in Table 5, column 3 (note that not all years can be included here as we are unable to separate the contribution type for all years).

Discussion and Pathway Analyses

We explore a series of pathways in which race/ethnicity could affect savings behavior. We explore several pathways put forth in the literature on the wealth gap in our data using simple mediation analyses. The pathway analyses assume that race is associated with a mediating variable, and that this mediating factor can explain a portion of the relationship between race and saving behavior. We implement these analyses in the spirit that attenuation potentially evidences a valid pathway, whereas a negative result (no effect) does not imply the absence of a pathway but rather, is more likely due to limitations in our data.

Family Size and Composition

For 56% percent of the sample, we know the number of children in the household under six years old for each year. This is an important cut-off because childcare is quite expensive until children enter school. If African

Americans and Hispanics have larger households, particularly with more young children, then the difference in participation and saving amounts could be a mechanical relation, where retirement savings is reduced to pay for necessary childcare. When we estimate the mediation model, Model 2, with number of children under age 6 in each year for each employee in our cohort, we find no substantial change in the magnitude of participation, contributions, balances, loans and withdrawals for Hispanics or African Americans (See Tables 6 and Table 7). Unfortunately, we do not have yearly information on spouses or cohabitation of domestic partners which also affects household size and composition.

Risk Aversion

While we do not have access to employees to ask standard risk aversion questions, we have a behavioral proxy for this construct. In particular, sometime in the period 2003-2006⁵, the employer-sponsored insurance options were changed for a subsample of employees in such a way that there was no default option, necessitating an active choice on the part of the employee. In addition, the company insurance options all include the same provider network, therefore there is no potential that insurance choices are being made to keep a particular provider or other unique aspect of the benefit unrelated to financial design. Given that insurance decisions expose workers to varying amounts of financial risk and the monetary differences between insurance options are often economically meaningful, choices on insurance could serve as a measure of risk aversion and are potentially more realistic than measures based on hypothetical lotteries often used in surveys. A similar measure has been used previously as a measure of relative risk aversion (Einav, Finkelstein et al. 2012). There are 5 plans/levels⁶, which we use as a proxy for an employee's risk aversion, and we include this time invariant variable as a mediating variable.

The addition of risk aversion substantially attenuates the differences in participation rates for African Americans and Hispanics relative to White. While African Americans still participate less, the magnitude of the coefficient on African American is essentially halved (see Tables 8 and Table 9) and Hispanics no longer participate at a lower rate than whites. This suggests that risk aversion mediates the decision to participate in the employer sponsored 401(k) plan. However, conditional on participation, African Americans still contribute 1.5% less. In addition, if we examine allocations, we find that the addition of this measure or relative risk aversion eliminates the significance of ethnic differences in choosing stock funds and the most risk adverse group contributes to stocks less (see Table 9). This measure is not a perfect proxy for pure risk aversion because it likely includes considerations of pre-existing health conditions and consumption tradeoffs. While the estimated gradients on this risk aversion measures are not perfectly monotonic, the patterns are largely inline with theory.

Financial Literacy

Financial literacy has been shown to substantially affect savings behavior. Moreover, there are substantial differences in financial literacy by education (Lusardi and Mitchell 2011). While we do not have a direct measure of financial literacy, we can distinguish whether workers were hourly or salaried employees. This employee type variable can serve as a crude proxy for financial literacy based on the idea that salaried workers generally have a college education while hourly workers do not (Worthington 2006, Lusardi and Mitchell 2007, Delavande, Rohwedder et al. 2008). Previously we include this variable as a control variable in model 1. Excluding this variable does lead to larger differences in the magnitudes of participation rates and savings as a portion of income, which suggests that employee status does mediate the relationship somewhat (analysis not shown, but available). However, it may be that much of the ethnic differences are driven by the differences in

⁵ The date the change occurred for each employee depends on the employee's plant.

⁶ The plans are increasing numerically in their coverage level. Hence 1 represents the plan which has the least coverage and is taken to represent the lowest level of risk aversion, whereas 5 represents the highest level of risk aversion.

racial composition of employee types. We therefore include an interaction term between employee status and race to investigate these relations further. We find that salaried African Americans are less likely to participate relative to salaried whites (OR = 0.30, 95% CI = [0.13, 0.67]), but Hispanic salaried workers participate at a similar rate as white salaried workers. Moreover, hourly African Americans and Hourly Hispanics participate less than hourly white workers (See Table 10).

Network Effects

In order to measure the effect of peer groups on 401(k) behavior, at each location we calculate the proportion of people in the same gender, ethnicity and employment type that participate in the 401(k) plans. We use this variable as proxy for how people within each workers peer groups participate in the 401(k) plans, and include this in the regression models as a mediating factor. We find the including this variable does mediate the relationship between race and savings behavior substantially (See Table 11). We interpret these results to suggest that peer groups may act as vehicle for information dissemination; however, we interpret this mediation analysis with caution because peer networks are assumed based on demographic and location characteristics alone.

Conclusions

In this paper we demonstrate that even within a single employer architecture, there is substantial difference in 401(k) savings behavior by employee ethnicity. We show both African American and Hispanic employees are less likely to participate in the 401(k) plans. Moreover, conditional on participation African Americans contribute a lower proportion of their income to their 401(k) plan on average. We also show that African Americans and Hispanics tend to draw down on their 401(k) balances more often through withdrawals and loans and this may cause lower average accumulations among these ethnicities over time. We also show that both African Americans and Hispanics tend to favor safer assets as opposed to Whites and this should have strong impact on the level of balances accumulated by these individuals and therefore overall wealth.

We undertake several supplementary analyses to better understand pathways that underlie some of the differences in savings behaviors across ethnicities. To explore this we conduct mediation analyses, using proxy variables for household composition, risk aversion, financial literacy, and peer networks. In each case the measurement of the construct of interest is imprecise relative to the underlying construct of interest. Nonetheless, we find suggestive evidence that both risk aversion and peer networks mediate some of the relationship between ethnicity and savings behavior.

There are other important factors that we would have liked to include in our analyses. For example, household income is an important variable for which we have no measure. If income shocks are distributed unevenly across groups and if such shocks affect savings behavior, then they are likely to mediate the observed relationship. Another variable of interest would be a measure of institutional distrust. If the legacy of contentious race relations has made certain ethnicities, such as African Americans and Hispanics, less likely to trust or invest in established/formal sector financial institutions, then that too could mediate the relationships observed. While these important constructs are missing in our mediation analyses, the primary results that there are substantial differences in savings behavior by race still hold. In future analyses we will link employees to households, allowing us to account for the impact of spousal income on savings behavior. This will allow us to address the issue of household income shocks. Regrettably we do not have an apparent data source to further explore psychological pathways, such as institutional distrust within the current data framework.

References

- Apgar, W. C. and A. Calder (2005). The Dual Mortgage Market: The Persistence of Discrimination in Mortgage Lending The Geography of Opportunity: Race and Housing Choice in Metropolitan America. Washington, D.C., Brookings Institution Press: 101-123.
- Argento, R., et al. (2015). "Early Withdrawals from Retirement Accounts During the Great Recession." Contemporary Economic Policy **33**(1): 1-16.
- Barsky, R. B., et al. (1997). "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey." Quarterly Journal of Economics **112**(2): 537-579.
- Basett, W. F., et al. (1998). "How workers use 401(k) plans: the participation, contribution, and withdrawal decisions." National Tax Journal **51**(2): 263-289.
- Beshears, J., et al. (2010). The Impact of 401 (k) Loans on Saving. NBER Working Papers, National Bureau of Economic Research.
- Borgo, M. D. (2013). "Ethnic and Racial Disparities in Saving Behavior." Institute for the Study of Labor. from http://www.iza.org/conference_files/SUMS_2013/dal_borgo_m5011.pdf.
- Bricker, J., et al. (2012). Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances. Federal Reserve Bulletin. **98**.
- Chetty, R., et al. (2012). Active vs Passive Decisions and Crowd-out In Retirement Savings Accounts: Evidence from Denmark. NBER Working Papers.
- Choi, J. J., et al. (2002). "Defined contribution pensions: plan rules, participant decisions, and the path of least resistance." Tax Policy and the Economy **16**: 68-113.
- Choudhury, S. (2001). "Racial and ethnic differences in wealth and asset choices." Soc Secur Bull **64**(4): 1-15.
- Clark, R. L. and S. Schieber (1998). Factors Affecting Participation Rates and Contribution Levels in 401(k) Plans. Living with Defined Contributions Pensions. O. S. S. Mitchell, Sylvester. Philadelphia, Pennsylvania, University of Pennsylvania Press: 69-97.
- Coates, T.-N. (2014). The Case for Reparations. The Atlantic. Washington, D.C., Atlantic Media Company: 1-62.
- Cole, S. and G. K. Shastry (2009). Smart Money: The Effect of Education, Cognitive Ability, and Financial Literacy on Financial Market Participation, Harvard Business School.
- Corbetti, C. and C. Hill (2014). The Simple Truth about the Gender Pay Gap. American Association of University Women Report, American Association of University Women.
- Delavande, A., et al. (2008). Preparation for Retirement, Financial Literacy and Cognitive Resources. Michigan Retirement Research Center Working Papers.
- Duflo, E. and E. Saez (2002). "Participation and investment decisions in a retirement plan: the influence of colleagues' choices." Journal of Public Economics **85**(1): 121-148.

EBRI (1994). Salary Reduction Plans and Individual Saving for Retirement. Employee Benefits Research Institute Research Brief. Washington, D.C. **155**.

Einav, L., et al. (2012). "How General Are Risk Preferences? Choices under Uncertainty in Different Domains." American Economic Review **102**(6): 2606–2638.

Engelhardt, G. V. and A. Kumar (2007). "Employer matching and 401(k) saving: Evidence from the health and retirement study." Journal of Public Economics **91**(10): 1920-1943.

Evans, W. N., et al. (1992). "Measuring Peer Group Effects: A Study of Teenage Behavior." Journal of Political Economy **100**(5): 966-991.

Gittleman, M. and E. N. Wolff (2004). "Racial Differences in Patterns of Wealth Accumulation." The Journal of Human Resources **39**(1): 193-227.

Gustman, A. L., et al. (2008). Do Workers Know Their Pension Plan Type? Comparing Workers' and Employers' Pension Information. Overcoming the Saving Slump. A. Lusardi, University of Chicago Press: 47-81.

Heinberg, A., et al. (2014). Five Steps to Planning Success. Experimental Evidence from U.S. Households. NBER Working Papers.

Hong, H., et al. (2005). "Social Interaction and Stock-Market Participation." The Journal of Finance **59**(1): 137-163.

Karlan, D. S. (2007). "Social Connections and Group Banking." The Economic Journal **117**(517): F52-F84.

Longhofer, S. D. a. P., Stephen R., (2004). "Self-Selection and Discrimination in Credit Markets " SSRN. from <http://ssrn.com/abstract=118088>

Love, D. A. (2007). "What can the Life-cycle Model Tell us about 401 (k) Contributions and Participation?" Journal of Pension Economics and Finance **6**(02): 147-185.

Lundberg, S. J. and J. Ward-Batts (2000). "Saving for retirement: Household bargaining and household net worth." Michigan Retirement Research Center Research Paper(2000-004).

Lusardi, A. and O. S. Mitchell (2007). "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." Journal Of Monetary Economics **54**(1): 205-224.

Lusardi, A. and O. S. Mitchell (2014). "The Economic Importance of Financial Literacy: Theory and Evidence." Journal of Economic Literature **52**(1): 5-44.

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

Mitchell, O. S. and J. S. Hastings (2011). How Financial Literacy and Impatience Shape Retirement Wealth and Investment Savings. NBER Working Papers.

Mitchell, O. S., et al. (2005). Turning Workers into Savers? Incentives, Liquidity, and Choice in 401(k) Plan Design. NBER Working Papers.

- Modrek, S., et al. (2015). "Psychological Well-being during the Great Recession: Changes in Mental Healthcare Utilization in an Occupational Cohort." American Journal of Public Health **105**(2): 304-310.
- Munell, A. H., et al. (2001). What Determines 401(k) Participation and Contributions. Social Security Bulletin. **64**: 64-75.
- Passel, J., et al. (2012). "Explaining why minority births now outnumber white births." from <http://pewsocialtrends.org/2012/05/17/explaining-why-minority-births-now-outnumber-white-births>.
- Pence, K. E. (2002). 401(k)s and Household Saving: New Evidence from the Survey of Consumer Finances, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.
- Pepke, L. E. and J. M. Poterba (1995). "Survey evidence on employer match rates and employee saving behavior in 401(k) plans." Economics Letters **49**: 313-317.
- Pollack, C. E., et al. (2013). "Do wealth disparities contribute to health disparities within racial/ethnic groups?" J Epidemiol Community Health **67**(5): 439-445.
- Rugh, J. S. and D. S. Massey (2010). "Racial Segregation and the American Foreclosure Crisis." American Sociological Review **75**(5): 629-651.
- Savage, C. (2011). Countrywide Will Settle a Bias Suit. New York Times. New York, New York.
- Savage, C. (2012). Wells Fargo Will Settle Mortgage Bias Charges. New York Times. New York, New York.
- Scholz, J. K. and A. Seshadri (2007). Children and household wealth. Michigan Retirement Research Center Working Papers.
- Shah Goda, G., et al. (2012). What Will My Account Really Be Worth? An Experiment on Exponential Growth Bias and Retirement Saving. NBER Working Papers.
- Shapiro, T., et al. (2013). The roots of the widening racial wealth gap: Explaining the black-white economic divide. . Institute on Assets and Social Policy Research and Policy Brief. Waltham, MA, Brandeis University.
- Taylor, P., et al. (2011). Wealth Gaps Rise to Record Highs Between Whites, Blacks and Hispanics. Pew Social & Demographic Trends. Pew Research Center. Washington, D.C, Pew Research Center, .
- VanDerhei, J. and C. Copeland (2001). "A Behavioral Model for Predicting Employee Contributions to 401(k) Plans: Preliminary Results." North American Actuarial Journal **5**(1): 80-94.
- VanDerhei, J., et al. (2009). 401(k) Plan Asset Allocation, Account Balances, and Loan Activity in 2008 Employee Benefit Research Institute Issue Brief.
- Webel, A. R., et al. (2010). "A Systematic Review of the Effectiveness of Peer-Based Interventions on Health-Related Behaviors in Adults." American Journal of Public Health **100**(2): 247-253.
- White, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." Econometrica(48): 817-830.
- Worthington, A. C. (2006). "Predicting Financial Literacy in Australia." Financial Services Review **15**(1): 59-79.

Figure 1. Pathways for Differences in 401(k) Behavior

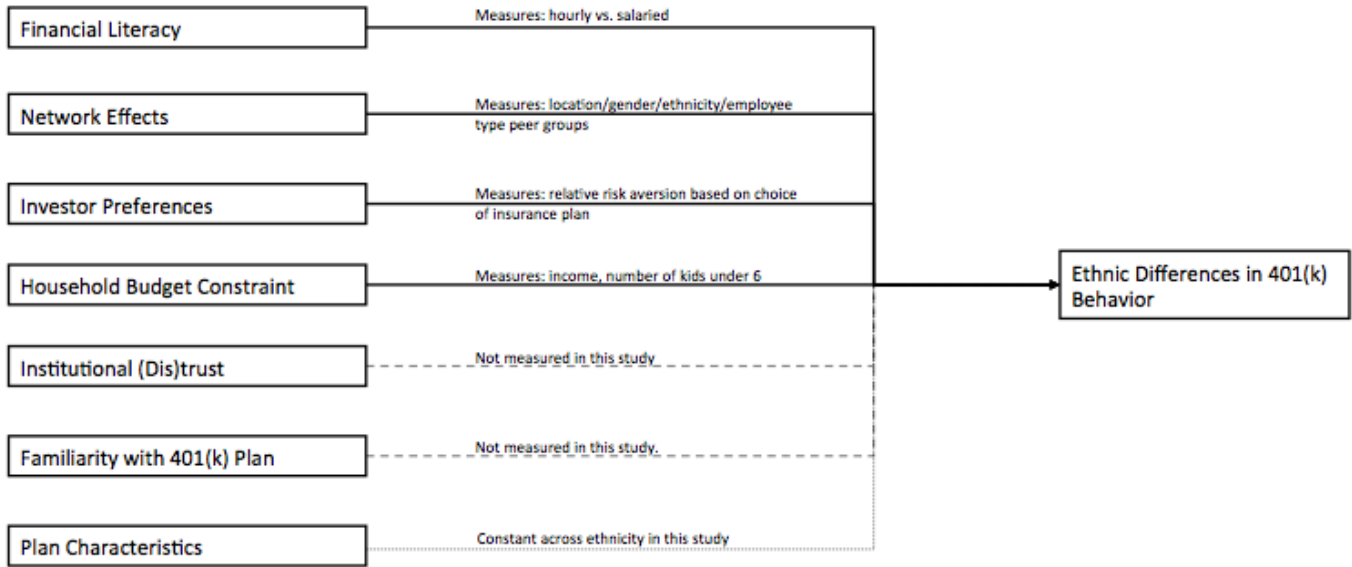


Figure 2. Participation Rate by Age and Race

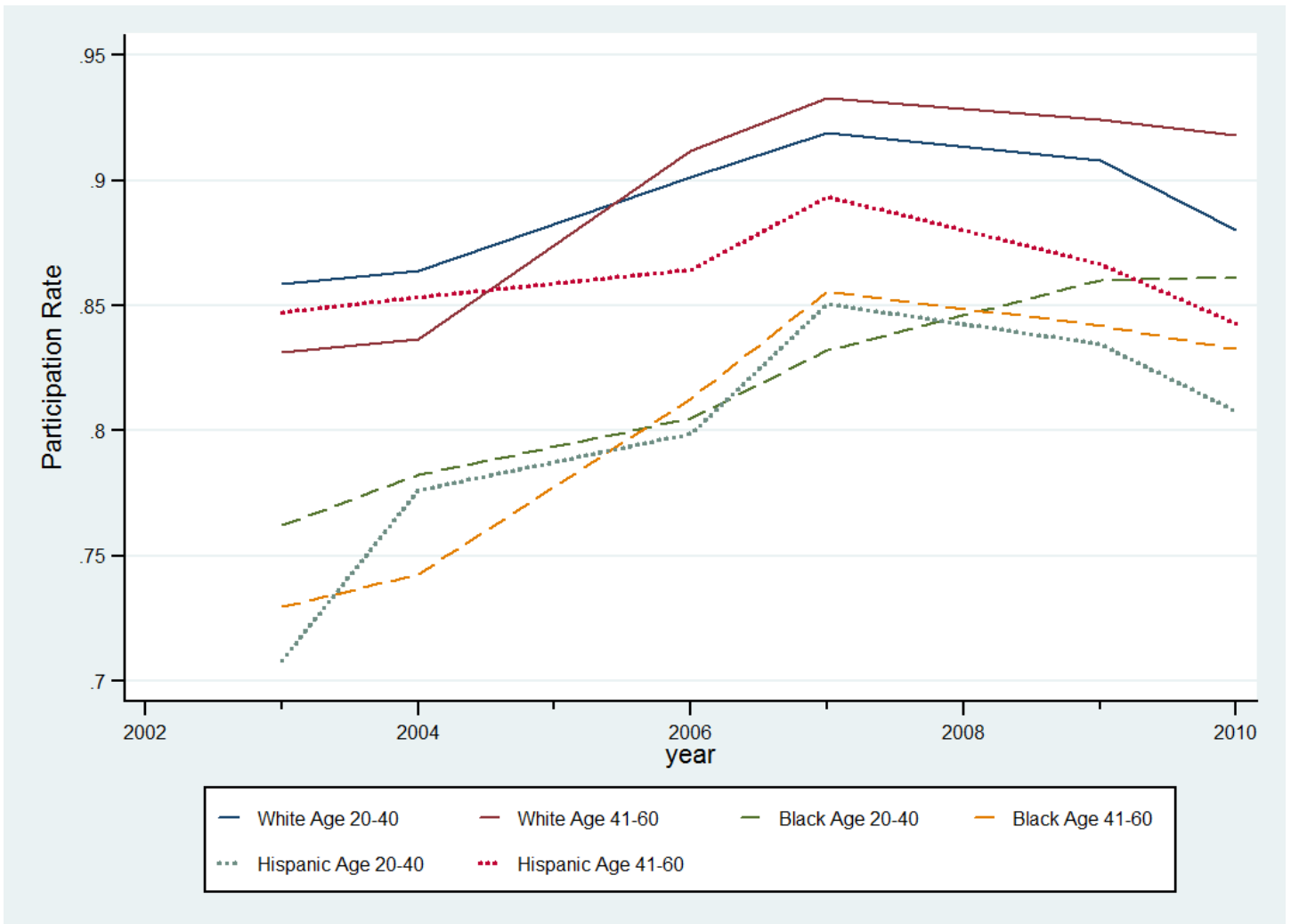


Figure 3. Median Contributions by Age and Race

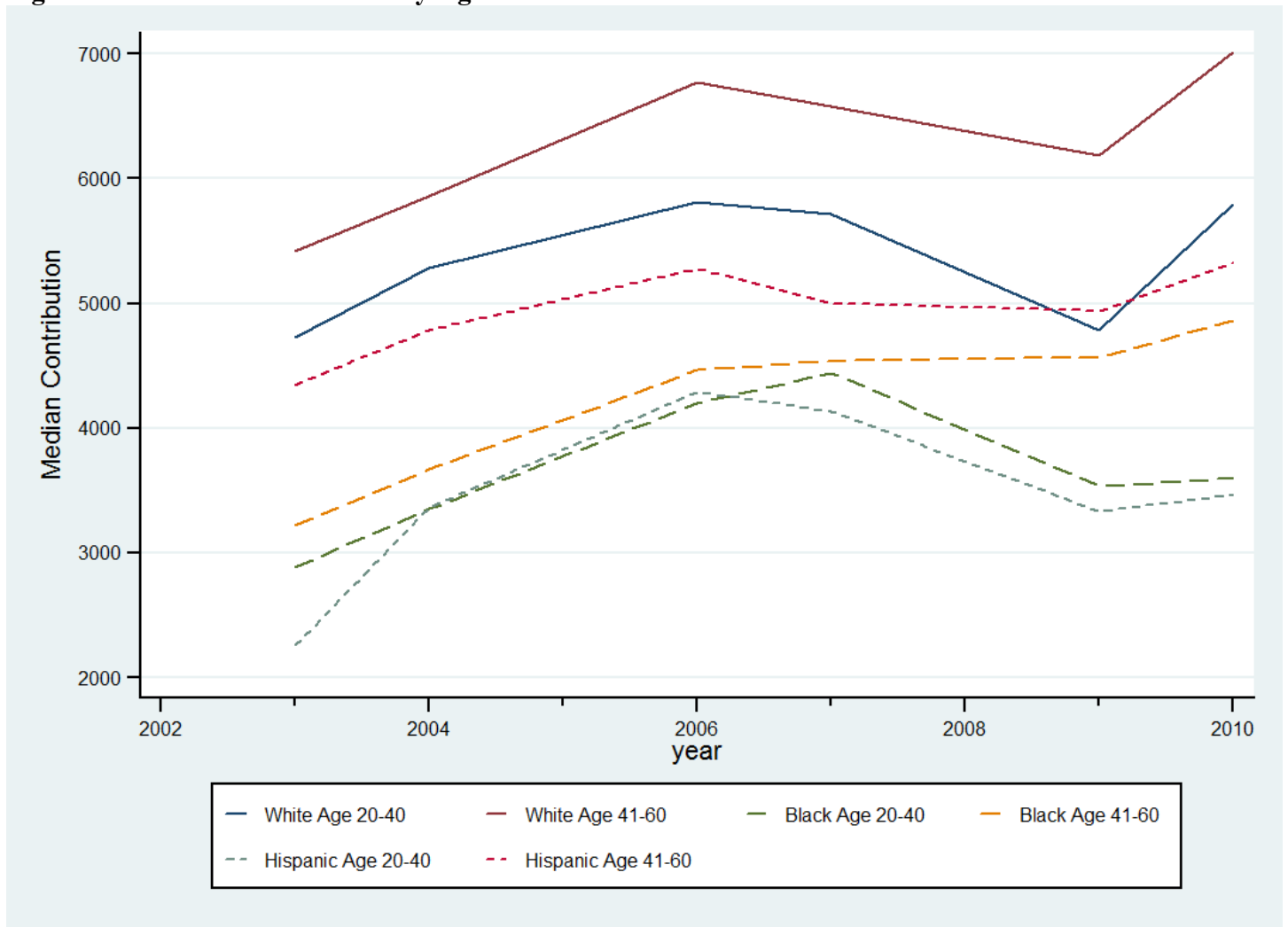


Figure 4. Median Balances by Age and Race

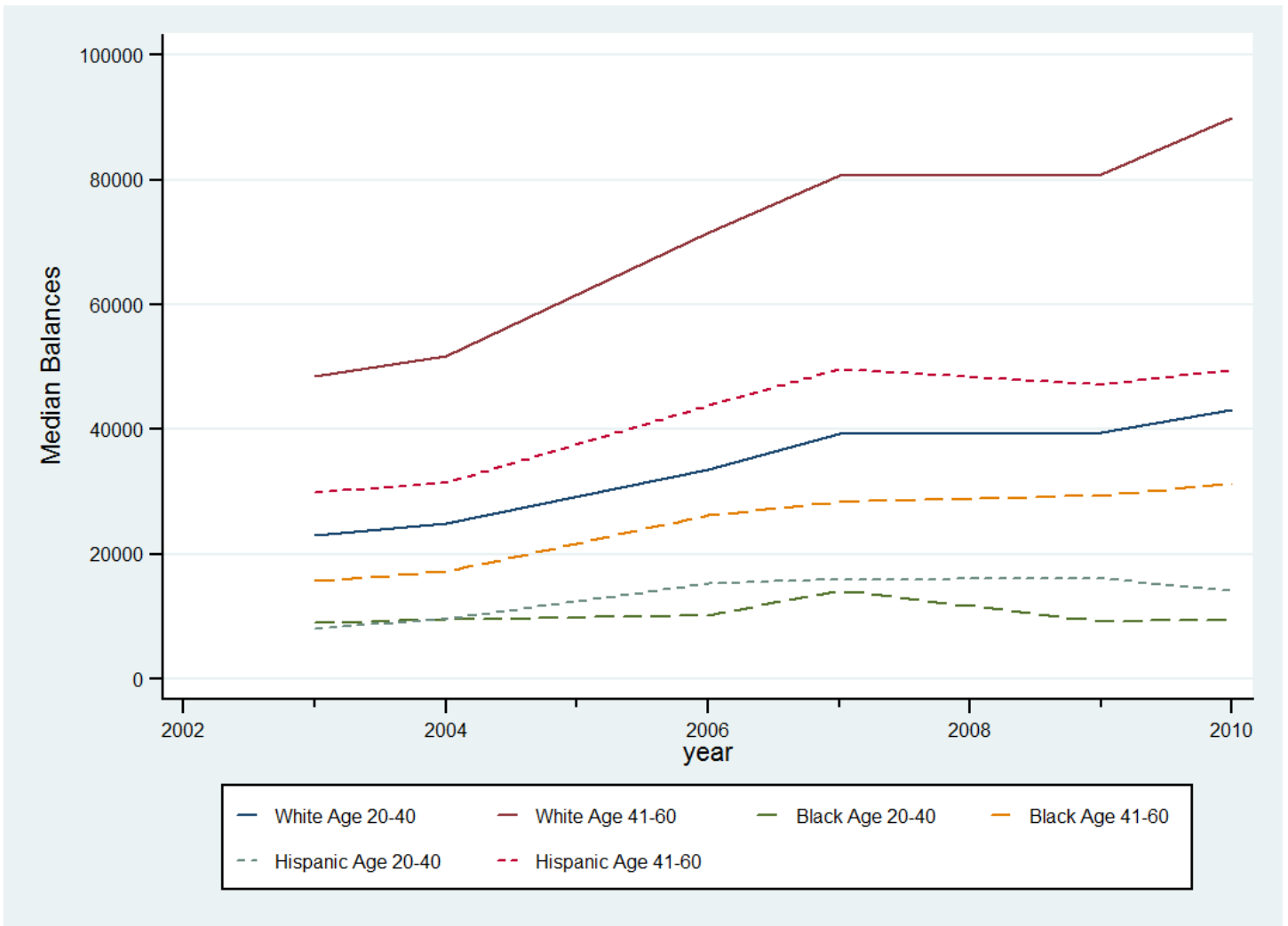


Figure 5. Withdrawal Rate by Age and Race

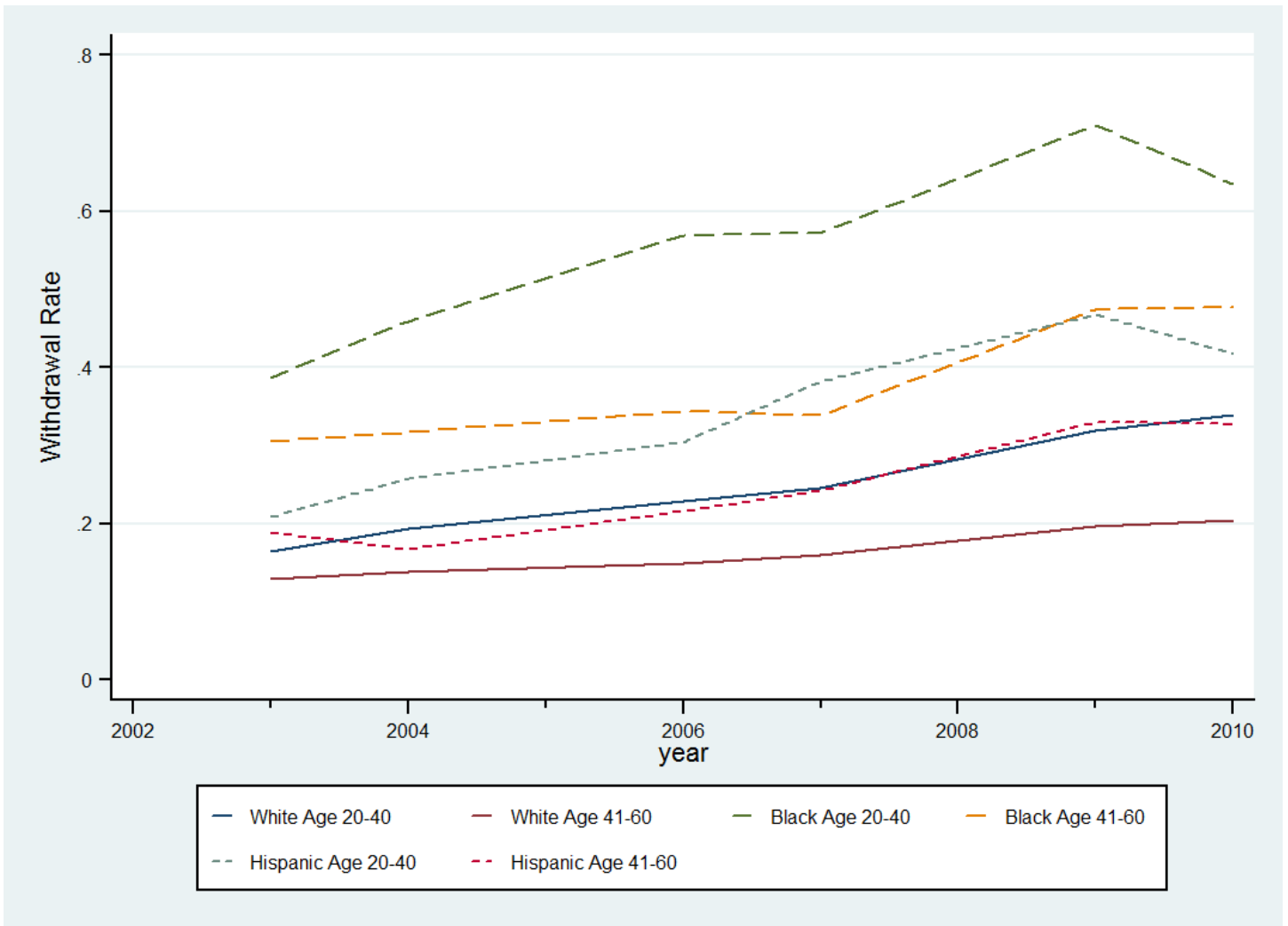


Figure 6. New Loan Rate by Age and Race

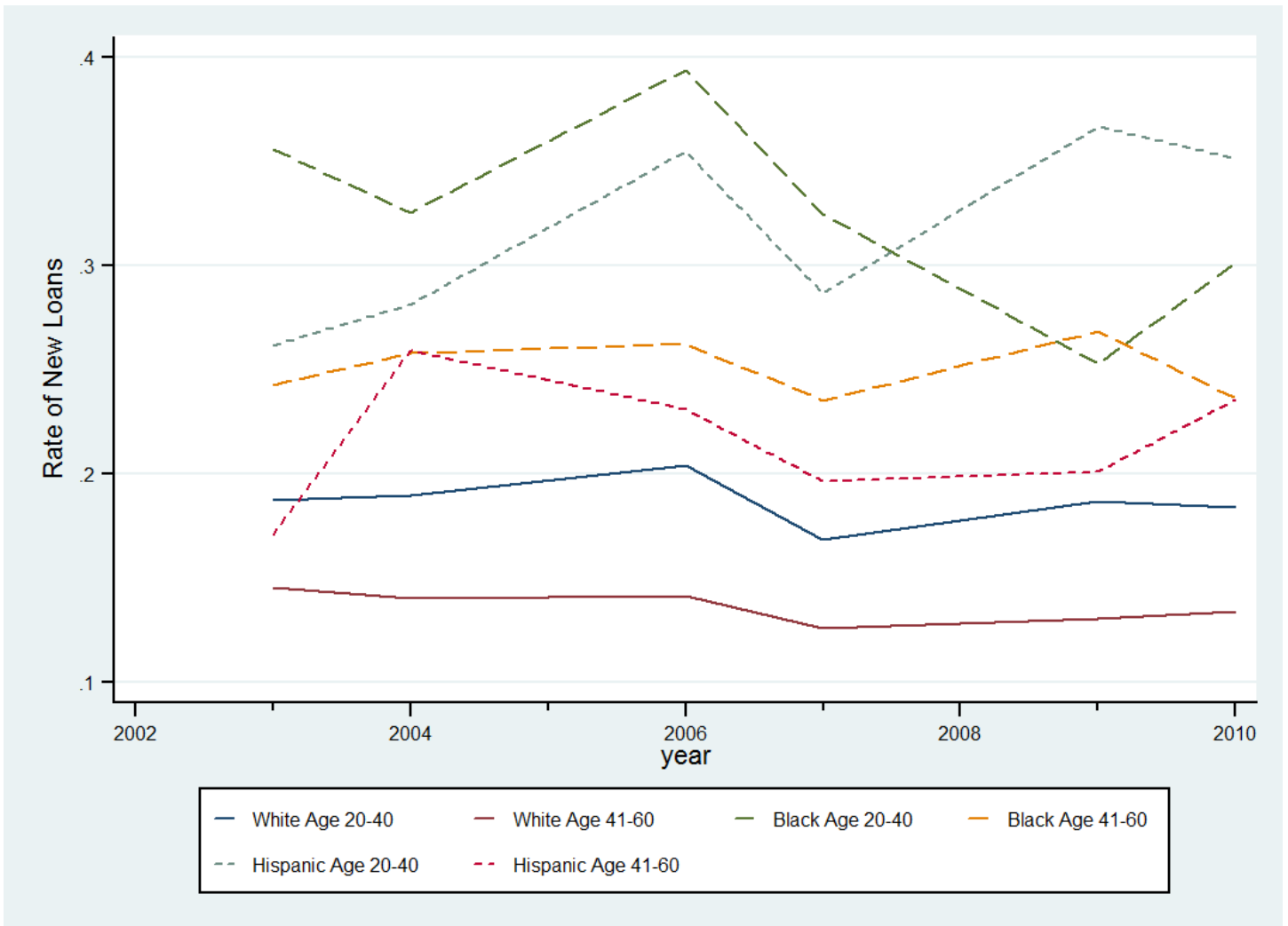


Table 1: Cohort Descriptive Statistics and Outcome Statistics

Panel A: Cohort Descriptive (N= 9,622)		
	Percent	
Gender		
Male	82.52%	
Birth Cohort		
Pre-1945	1.53%	
1946-1950	10.35%	
1951-1955	24.12%	
1956-1960	26.08%	
1961-1965	17.48%	
1966-1970	11.53%	
1971-1975	6.36%	
Post-1976	2.56%	
Ethnicity		
White	83.09%	
Black	8.29%	
Hispanic	5.79%	
Asian	2.04%	
Other	0.79%	
Risk Scores		
0-1	64.11%	
1-2	26.17%	
2-3	6.62%	
3-4	2.10%	
4-6	1.00%	
Employee Type		
Hourly	76.43%	
Annual Pay (Mean and SD)	\$56,283	\$19,181
Panel B: 401K outcomes (N=50822)		
401K outcomes		
Participated (N=57,732)	88.03%	
Contributions (Mean and SD)	\$7,060	\$5,007
Withdrew	19.97%	
Got Loans	15.57%	
Contributed to Stock Funds	62.83%	
Contributed to Bond Funds	24.88%	
Contributed to Money Market Funds	56.83%	

Table 2: Participation, Contributions and Balances

	Participation (OR)^{&}	Log Contributions	Percent Income Contributed	Log Balances
<u>Race (White Omitted)</u>				
African American	0.264*** (0.0533)	-0.193*** (0.0251)	-1.867*** (0.196)	-0.610*** (0.0455)
Hispanic	0.361*** (0.0840)	-0.0312 (0.0298)	-0.342 (0.266)	-0.187*** (0.0517)
Other	1.078 (0.306)	0.140*** (0.0429)	1.650*** (0.421)	-0.00432 (0.0734)
<u>Risk (Risk Score 0-1 Omitted)</u>				
Risk Score 1-6	0.965 (0.0795)	-0.0180*** (0.00583)	-0.0469 (0.0481)	-0.00913 (0.00574)
<u>Gender (Female Omitted)</u>				
Male	0.965 (0.0795)	-0.0180*** (0.00583)	-0.0469 (0.0481)	-0.00913 (0.00574)
<u>Employee Type (Hourly Omitted)</u>				
Salaried	3.904*** (0.577)	0.401*** (0.0160)	3.397*** (0.163)	0.898*** (0.0269)
<u>Constant</u>	181.1*** (106.2)	7.508*** (0.0676)	10.62*** (0.507)	8.817*** (0.112)
Birth Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	48,508	43,882	43,882	43,795
Number of Unique Workers	9,066	8,521	8,521	8,516

Controls: Yearly Salary, Country-Level Unemployment

[&]First analysis is a logistic regression, where the dependent variable is a binary variable that takes on the value 1 if a contribution was made to any of his/her 401(k) savings accounts (by the individual or by the company). The coefficients are displayed in odds ratio format. The remaining models use linear regressions.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 3: Withdrawals and Loans

	Withdrawals (OR)	Loans (OR)
<u>Ethnicity (White Omitted)</u>		
African American	8.207*** (1.188)	2.187*** (0.146)
Hispanic	1.835*** (0.339)	1.713*** (0.141)
Other	1.142 (0.289)	1.136 (0.144)
<u>Risk (Risk Score 0-1 Omitted)</u>		
Risk Score 1-6	1.252*** (0.0638)	1.015 (0.0370)
<u>Gender (Female Omitted)</u>		
Male	0.581*** (0.0664)	0.791*** (0.0440)
<u>Employee Type (Hourly Omitted)</u>		
Salaried	0.540*** (0.0590)	0.575*** (0.0324)
<u>Constant</u>	0.888 (0.379)	0.216*** (0.0534)
Birth Cohort FE	Yes	Yes
Year FE	Yes	Yes
Location FE	Yes	Yes
Observations	43,882	43,882
Number of Unique Workers	8,521	8,521

Controls: Salary, Country-Level Unemployment

Both analyses are logistic regressions, where the dependent variable is binary and takes on the value 1 if a withdrawal or loan was made by the individual in the given year. The coefficients are displayed in odds ratio format.

*Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$*

Table 4: Allocation Mix Models

	Contributed to Stock Funds (OR)	Contributed to Bond Funds (OR)	Contributed to Money Market Funds (OR)
<u>Ethnicity (White Omitted)</u>			
African American	0.426*** (0.109)	0.820 (0.142)	2.849*** (0.572)
Hispanic	0.489** (0.152)	0.509*** (0.107)	1.954*** (0.477)
Other	0.667 (0.263)	0.677 (0.194)	2.399*** (0.729)
<u>Risk (Risk Score 0-1 Omitted)</u>			
Risk Score 1-6	1.028 (0.0590)	1.072 (0.0556)	1.133** (0.0573)
<u>Gender (Female Omitted)</u>			
Male	0.628*** (0.0975)	0.621*** (0.0790)	0.689*** (0.0964)
<u>Employee Type (Hourly Omitted)</u>			
Salaried	5.541*** (0.788)	3.438*** (0.409)	0.741** (0.0933)
<u>Constant</u>	38.20*** (23.21)	0.217*** (0.105)	3.917*** (1.974)
Birth Cohort FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Observations	43,882	43,882	43,882
Number of Unique Workers	8,521	8,521	8,521

Controls: Salary, Country-Level Unemployment

All analyses are logistic regressions, where the dependent variable is a binary variable that takes on the value 1 if a 401(k) investment was made into a stock fund, bond fund or money market fund in the given year. The coefficients are displayed in odds ratio format.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 5: Contribution Type Robustness

	Participation (OR)&	Log Contributions	Log After Tax Contribution
<u>Ethnicity (White Omitted)</u>			
African American	0.237*** (0.0495)	-0.170*** (0.0255)	-0.186*** (0.0257)
Hispanic	0.336*** (0.0803)	-0.00827 (0.0304)	-0.0105 (0.0305)
Other	1.111 (0.307)	0.171*** (0.0439)	0.142*** (0.0454)
<u>Risk (Risk Score 0-1 Omitted)</u>			
Risk Score 1-6	0.936 (0.0905)	-0.0122 (0.00711)	-0.0105 (0.00708)
<u>Gender (Female Omitted)</u>			
Male	1.137 (0.167)	0.102*** (0.0178)	0.103*** (0.0179)
<u>Employee Type (Hourly Omitted)</u>			
Salaried	2.582*** (0.376)	0.233*** (0.0168)	0.218*** (0.0169)
<u>Constant</u>	172.4*** (108.6)	6.877*** (0.0643)	6.871*** (0.0648)
Birth Cohort FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Observations	40,832	37,033	36,663
Number of Unique Workers	9,598	9,033	8,968

Controls: Salary, Country-Level Unemployment

& First analysis is a logistic regression, where the dependent variable is a binary variable that takes on the value 1 if a contribution was made by the individual to any of his/her 401(k) savings accounts. The coefficients are displayed in odds ratio format. The remaining two models use linear regressions.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 6: Family Size - Participation, Contributions and Balances

	Participation (OR)^{&}	Log Contributions	Percent Income Contributed	Log Balances
<u>Ethnicity (White Omitted)</u>				
African American	0.234*** (0.0571)	-0.176*** (0.0299)	-1.616*** (0.224)	-0.660*** (0.0545)
Hispanic	0.315*** (0.0864)	-0.0145 (0.0344)	-0.0931 (0.293)	-0.180*** (0.0623)
Other	0.991 (0.323)	0.114** (0.0521)	1.399*** (0.473)	-0.0571 (0.0856)
<u>Number of Kids Under 6 (0 Kids Omitted)</u>				
1 Kid	1.092 (0.170)	0.00642 (0.0123)	0.176* (0.0983)	0.0388*** (0.0150)
2 Kids	1.094 (0.276)	0.0470** (0.0206)	0.582*** (0.163)	0.0530** (0.0247)
3 Kids	0.639 (0.338)	0.0632 (0.0476)	0.734** (0.337)	0.0311 (0.0584)
4 Kids	6.495 (17.04)	-0.188 (0.244)	0.000776 (0.905)	-0.0334 (0.208)
<u>Risk (Risk Score 0-1 Omitted)</u>				
Risk Score 1-6	0.860 (0.0848)	-0.0111 (0.00655)	-0.0539 (0.0517)	-0.0119 (0.00649)
<u>Gender (Female Omitted)</u>				
Male	1.579** (0.311)	0.0710*** (0.0243)	0.582*** (0.216)	0.269*** (0.0451)
<u>Employee Type (Hourly Omitted)</u>				
Salaried	3.693*** (0.641)	0.359*** (0.0190)	3.191*** (0.176)	0.866*** (0.0320)
<u>Constant</u>	100.1*** (73.94)	7.466*** (0.0822)	10.60*** (0.579)	8.723*** (0.135)
Birth Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	39,480	35,944	35,944	35,877
Number of Unique Workers	6,292	5,961	5,961	5,959

Controls: Salary, Country-Level Unemployment

& First analysis is a logistic regression, where the dependent variable is a binary variable that takes on the value 1 if a contribution was made to any of his/her 401(k) savings accounts (by the individual or by the company). The coefficients are displayed in odds ratio format. The remaining models use linear regressions.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 7: Family Size - Withdrawals and Loans

	Withdrawals (OR)	Loans (OR)
<u>Ethnicity (White Omitted)</u>		
African American	7.624*** (1.296)	2.200*** (0.171)
Hispanic	1.628** (0.348)	1.710*** (0.162)
Other	1.050 (0.295)	1.203 (0.172)
<u>Number of Kids Under 6 (0 Kids Omitted)</u>		
1 Kid	0.896 (0.0826)	0.938 (0.0541)
2 Kids	0.924 (0.140)	0.799** (0.0728)
3 Kids	0.682 (0.261)	1.190 (0.259)
4 Kids	2.007 (2.005)	0.589 (0.382)
<u>Risk (Risk Score 0-1 Omitted)</u>		
Risk Score 1-6	1.283*** (0.0770)	0.976 (0.0420)
<u>Gender (Female Omitted)</u>		
Male	0.564*** (0.0849)	0.798*** (0.0568)
<u>Employee Type (Hourly Omitted)</u>		
Salaried	0.553*** (0.0682)	0.574*** (0.0364)
<u>Constant</u>	1.034 (0.525)	0.212*** (0.0606)
Birth Cohort FE	Yes	Yes
Year FE	Yes	Yes
Location FE	Yes	Yes
Observations	31,160	31,160
Number of Unique Workers	5,960	5,960

Controls: Salary, Country-Level Unemployment

Both analyses are logistic regressions, where the dependent variable is binary and takes on the value 1 if a withdrawal or loan was made by the individual in the given year. The coefficients are displayed in odds ratio format.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 8: Risk Aversion

	Participation (OR) &	Log Contributions	Percent Income Contributed	Log Balances
<u>Ethnicity (White Omitted)</u>				
African American	0.418*** (0.112)	-0.137*** (0.0356)	-1.468*** (0.309)	-0.513*** (0.0674)
Hispanic	0.568 (0.183)	-0.0659 (0.0456)	-0.557 (0.420)	-0.181** (0.0804)
Other	1.552 (0.793)	0.160** (0.0679)	2.281*** (0.817)	0.114 (0.104)
<u>Risk Aversion (Level 1 Omitted)</u>				
Risk Aversion Level 2	0.486 (0.467)	-0.190** (0.0882)	-1.936** (0.822)	-0.111 (0.155)
Risk Aversion Level 3	0.744 (0.744)	-0.181*** (0.0649)	-2.574*** (0.673)	-0.151 (0.123)
Risk Aversion Level 4	0.176*** (0.103)	-0.217*** (0.0321)	-2.537*** (0.382)	-0.365*** (0.0595)
Risk Aversion Level 5	0.355 (0.197)	-0.187*** (0.0282)	-2.123*** (0.349)	-0.254*** (0.0523)
<u>Risk (Risk Score 0-1 Omitted)</u>				
Risk Score 1-6	0.966 (0.123)	-0.00139 (0.00906)	0.0465 (0.0821)	-0.0135 (0.00892)
<u>Gender (Female Omitted)</u>				
Male	0.792 (0.188)	0.146*** (0.0260)	0.849*** (0.269)	0.300*** (0.0477)
<u>Employee Type (Hourly Omitted)</u>				
Salaried	4.158*** (1.073)	0.338*** (0.0285)	2.896*** (0.279)	0.863*** (0.0476)
Constant	425.0*** (439.6)	7.870*** (0.0980)	12.50*** (0.873)	8.864*** (0.175)
Birth Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	19,347	17,561	17,561	17,528
Number of Unique Workers	3,289	3,145	3,145	3,144

Controls: Salary, Country-Level Unemployment

& First analysis is a logistic regression, where the dependent variable is a binary variable that takes on the value 1 if a contribution was made to any of his/her 401(k) savings accounts (by the individual or by the company). The coefficients are displayed in odds ratio format. The remaining three models use linear regressions.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 9: Risk Aversion – Allocation Mix

	Contributed to Stock Funds (OR)	Contributed to Bond Funds (OR)	Contributed to Money Funds (OR)
<u>Ethnicity (White Omitted)</u>			
African American	0.760 (0.231)	0.660 (0.165)	3.061*** (0.839)
Hispanic	0.762 (0.289)	0.422*** (0.128)	1.699 (0.565)
Other	2.180 (1.120)	1.047 (0.458)	1.075 (0.490)
<u>Risk Aversion (Level 1 Omitted)</u>			
Risk Aversion Level 2	0.241 (0.183)	0.169** (0.124)	8.552*** (5.995)
Risk Aversion Level 3	1.354 (0.693)	1.429 (0.677)	1.432 (0.705)
Risk Aversion Level 4	0.888 (0.233)	1.046 (0.244)	1.178 (0.301)
Risk Aversion Level 5	0.481*** (0.111)	1.138 (0.235)	1.272 (0.284)
<u>Risk (Risk Score 0-1 Omitted)</u>			
Risk Score 1-6	1.052 (0.0976)	1.099 (0.0818)	1.210** (0.0961)
<u>Gender (Female Omitted)</u>			
Male	0.771 (0.155)	1.060 (0.183)	0.536*** (0.100)
<u>Employee Type (Hourly Omitted)</u>			
Salaried	1.660** (0.379)	2.271*** (0.437)	0.873 (0.181)
<u>Constant</u>	55.80*** (45.36)	0.387 (0.264)	2.339 (1.666)
Birth Cohort FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Observations	17,561	17,561	17,561
Number of Unique Workers	3,145	3,145	3,145

Controls: Salary, Country-Level Unemployment

All analyses are logistic regressions, where the dependent variable is a binary variable that takes on the value 1 if a 401(k) investment was made into a stock fund, bond fund or money market fund in the given year. The coefficients are displayed in odds ratio format.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 10: Financial Sophistication

	Participation (OR)^{&}	Log Contributions	Percent Income Contributed	Log Balances
<u>Ethnicity-Employee Type Interactions (Salaried-White Omitted)</u>				
Hourly-African American	0.0716*** (0.0179)	-0.587*** (0.0288)	-5.167*** (0.247)	-1.496*** (0.0522)
Hourly-Hispanic	0.0893*** (0.0247)	-0.431*** (0.0332)	-3.759*** (0.298)	-1.084*** (0.0581)
Hourly-Other	0.233*** (0.0758)	-0.269*** (0.0500)	-1.955*** (0.421)	-0.913*** (0.0844)
Hourly-White	0.287*** (0.0451)	-0.408*** (0.0170)	-3.425*** (0.174)	-0.910*** (0.0285)
Salaried-African American	0.297*** (0.130)	-0.272*** (0.0634)	-2.643*** (0.524)	-0.758*** (0.118)
Salaried-Hispanic	0.772 (0.368)	-0.0681 (0.0746)	-0.371 (0.753)	-0.253** (0.125)
Salaried-Other	6.954 (7.156)	0.150 (0.0803)	2.398 (1.247)	-0.00154 (0.169)
<u>Risk (Risk Score 0-1 Omitted)</u>				
Risk Score 1-6	0.965 (0.0800)	-0.0180*** (0.00583)	-0.0471 (0.0481)	-0.00913 (0.00574)
<u>Gender (Female Omitted)</u>				
Male	1.066 (0.154)	0.0954*** (0.0177)	0.569*** (0.169)	0.258*** (0.0311)
<u>Constant</u>	717.8*** (438.9)	7.912*** (0.0683)	14.02*** (0.520)	9.720*** (0.113)
Birth Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	48,508	43,882	43,882	43,795
Number of Unique Workers	9,066	8,521	8,521	8,516

Controls: Salary, Country-Level Unemployment

& First analysis is a logistic regression, where the dependent variable is a binary variable that takes on the value 1 if a contribution was made to any of his/her 401(k) savings accounts (by the individual or by the company). The coefficients are displayed in odds ratio format. The remaining three models use linear regressions.

*Robust standard errors in parentheses *** p<0.01, ** p<0.05*

Table 11: Network Effects – Average Network Participation

	Participation (OR)	Contributed to Stock Funds (OR)	Contributed to Bond Funds (OR)	Contributed to Money Funds (OR)
<u>Ethnicity (White Omitted)</u>				
African American	0.470*** (0.0930)	0.473*** (0.113)	0.710** (0.122)	2.241*** (0.471)
Hispanic	0.584** (0.132)	0.563** (0.163)	0.528*** (0.110)	1.687** (0.408)
Other	1.020 (0.297)	0.733 (0.262)	0.828 (0.227)	2.153*** (0.627)
<u>Risk (Risk Score 0-1 Omitted)</u>				
Risk Score 1-6	0.977 (0.0786)	1.027 (0.0588)	1.072 (0.0554)	1.132** (0.0571)
<u>Gender (Female Omitted)</u>				
Male	- 0.913 (0.138)	- 0.687** (0.103)	- 0.759** (0.0962)	5.686* 0.779* (0.112)
<u>Employee Type (Hourly Omitted)</u>				
Salaried	1.312 (0.291)	3.539*** (0.541)	3.199*** (0.371)	0.921 (0.128)
<u>Constant</u>	0.109* (0.126)	1.265 (1.066)	0.0427*** (0.0223)	0.475 (0.392)
Birth Cohort FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	55,723	50,478	50,478	50,478
Number of Unique Workers	9,066	8,522	8,522	8,522

Controls: Salary, Country-Level Unemployment

*All analyses from logistic regression model, where the dependent variable is a binary variable. All coefficients are displayed in odds ratio format. Robust standard errors in parentheses *** p<0.01, ** p<0.05*