Strength in Differences: How Racial Integration Shapes Household Financial Decision-Making

Melina Murren Vosse *

Abstract

Using proprietary geocoded data from the Panel Study of Income Dynamics, I examine whether local racial integration influences financial decision-making. I find that individuals residing in racially integrated communities are more likely to invest in public equity markets after accounting for individual and county-level differences. Exploiting within-individual variation from relocation as well as using Great Migration population shocks as an instrument, I demonstrate that integration has a causal effect on participation. Evidence suggests that racial integration improves local information quality, lowering informational barriers to participation. Moreover, this informational advantage enables integrated investors to achieve superior risk-adjusted performance on their local portfolios.

Keywords: Household Finance, information environment, social structure, behavioral finance.

JEL classification: D14, D63, D83, D85, D91, G41, G51.

*Please address all correspondence to Melina Vosse, University of Miami, Miami Herbert Business School, email:mmurren@bus.miami.edu. I am thankful for insightful comments from William Bazley, Vidhi Chhaochharia, Stefanos Delikouras, George Korniotis, Alok Kumar, Ville Rantala and seminar participants at the University of Miami.

I. Introduction

Participation in financial markets is one of the primary channels through which households can accumulate wealth. Despite historically high returns in the stock market, the majority of the population forgoes this opportunity for wealth accumulation by not participating. Specifically, as of 2019, approximately 24% of households indicate owning publicly traded equity directly. While the literature has set forth numerous explanations for limited participation,[1] low participation rates remain a puzzle.

When breaking down participation rates by demographic groups, 29% of White Americans report owning equity, whereas this figure is 17% among minorities. Most concerning, however, is that the rate of participation is only 9% when restricting the minority group to include only Black Americans. This racial gap in asset ownership plays an important role in understanding the barriers to wealth accumulation for the minority population and is yet to be fully explained in the literature.

I propose that characteristics of the local environment can help explain the persistent gap in participation. In this paper, I isolate racial residential integration as a first pass at understanding whether local factors matter for financial decision-making. Specifically, I posit that racially integrated communities improve local information quality by enabling information transmission across friendship networks, effectively reducing informational barriers to participation. While social structures such as segregation have been examined in the field of economics and sociology, this concept has not yet been explored in the specific context of household financial decision-making.

[1] see Campbell (2006), Guiso, Sapienza, and Zingales (2018), Hong, Kubik, and Stein (2004), Ivković and Weisbenner (2005), Bonaparte, Kumar, and Page (2017), Bonaparte and Kumar (2013)Malmendier and Nagel (2011)

Using a detailed data set providing information on household location at the neighborhood level from the Panel Study of Income Dynamics (PSID), I find that individuals residing in integrated counties are approximately 4% more likely to invest in public equity markets relative to those residing in segregated counties. This magnitude is comparable to other factors of participation found in recent studies.[2] Additionally, this can be seen as a lower bound for local effects since racial integration is only one component within the complex structure of local environments.

Identifying a link between residential racial integration and the participation decision is empirically challenging. In the baseline specification I control for known determinants of individual financial decisions such as age, gender, education, income and wealth. Moreover, I ensure that the results are not driven by local socioeconomic conditions that may be correlated with diversity and an individual's choice to participate. All specifications, therefore, include controls to absorb irrelevant variation related to socioeconomic conditions such as median home value, urbanicity, social capital and county-level participation rates. Finally, I include a range of controls related to local publicly traded firms following Ivković and Weisbenner (2005) to ensure that the results are not biased by local financial market conditions.

I also employ two additional methods of identification to establish a causal relation between racial integration and participation. First, a concern might be that this result is driven by individuals who have a higher propensity to invest (based on education, income, etc.) systematically self-selecting into integrated communities. To address this criticism, I take advantage of the panel structure of the data set to test whether the same individual changes their behavior after moving from a segregated area to an integrated area. Consis-

[2]For example, Ivković and Weisbenner (2005) Giannetti and Wang (2016),Addoum, Delikouras, Ke, and Korniotis (2018)

tent with the baseline results above, I find that previously non-participating individuals who move from segregated to integrated communities are significantly more likely to participate after the move, with the effect strengthening over time. These findings suggest that the baseline results are not driven by self-selection of participating households into integrated areas and that household behavior is at least partially dependent on characteristics of the local environment.

Second, there remains a concern that the results may be driven by some omitted variable within a community that is correlated with participation and racial integration. To address this issue, I implement an instrumental variables specification where integration is instrumented for with the predicted Black population shock to northern and western commuting zones during the Great Migration of 1940-1970 (Derenoncourt (2019)). This instrument captures only the variation relevant to the spatial integration of individuals without directly affecting the household's participation decision. The results from this specification support the baseline results. Moreover, employing only the variation relevant to racial strengthens the economic magnitude of the effect.

The next set of tests examines whether integration has a disproportionate influence on the participation decision of minority or majority households. One hypothesis is that, in integrated communities, information relevant to financial investment is transferred from the majority (White) stock owning population to the minority (Black) population. If this were the case, the presence of diversity should not significantly impact the participation decision of an individual in the majority group. Alternatively, local integration may generate a systemic effect on the community and provide a pareto improvement in financial outcomes for all community members. Results from a specification modeling the marginal effect of integration

for White and Black households supports the latter hypothesis. This key result emphasizes the systematic importance of community-level factors in financial decision making.

Having established that racial integration plays a significant role in the participation decision, I move on to identify how integration affects financial decision-making. Work on network structures has found that greater network diversity improves the availability of information (Eagle, Macy, and Claxton (2010), Granovetter and Soong (1983)). The intuitive argument is that heterogeneity in social ties increases the likelihood of network overlap. This network overlap enables information to spread more easily across groups. Therefore, I conjecture that more diverse social ties in integrated environments improves the likelihood that individuals are exposed to information relevant to investing.

If integration improves the local information environment, investors should be able to make more informed financial decisions and achieve superior portfolio performance. I test this conjecture using a detailed data set that provides the monthly portfolio positions as well as the individual trades of investors in a large discount brokerage. I find that investors residing in integrated counties achieve higher risk-adjusted performance on the local portion of their portfolios, indicating that they have access to higher quality information. These investors recognize their informational advantage and display overconfidence and riskier trading behavior such as over-weighting local stocks, holding less diversified portfolios and turning over their portfolios more frequently. While integrated investors display more biased behavior, these seemingly irrational investment decisions end up being rewarded with superior risk-adjusted local performance.

Evidence from the trading behavior and portfolio decisions of investors identifies integration's effect on the local information environment as the most likely mechanism. While alternative hypotheses are explored, the evidence of lower diversification, higher turnover,

and superior performance in combination with higher participation rates is difficult to reconcile with alternative mechanisms such as financial literacy or wealth. Additionally, other factors that may coincide with higher participation rates and overconfidence—such as gender and income—are controlled for in the regressions.

Among the myriad factors that have been found to affect stock market participation, the role of the local social structures has been largely ignored. Prior literature has found significant effects of sociability and word-of-mouth communication (Hong, Kubik, and Stein (2004), Brown, Ivković, Smith, and Weisbenner (2008)), however these papers do not address whether the structure of the community itself can influence the behavior of individuals. These studies also posit that effects of sociability are driven by interaction with other stock market participants, insinuating that the benefit of sociability is achieved primarily through the exchange of direct information related to investing. Unexplored, however, is understanding whether other characteristics of communities —seemingly unrelated to financial investment—can influence financial decisions. In recognizing community structure as an additional dimension to participation, we can obtain a better understanding for what drives individuals to participate in financial markets.

This paper contributes to the literature by identifying community structure, namely the degree of racial integration, as an important determinant for the stock market participation decision. Broadly, this paper adds to the literature exploring determinants of household financial decision making (Campbell (2006), Guiso, Sapienza, and Zingales (2018), Malmendier and Nagel (2011), Bonaparte and Kumar (2013), Bonaparte, Kumar, and Page (2017), Behrman, Mitchell, Soo, and Bravo (2012), Ke (2021)). More closely related are the papers establishing the role of awareness and sociability on the participation decision (Guiso and Jappelli (2005)Hong, Kubik, and Stein (2004), Brown, Ivković, Smith, and Weisbenner

(2008)). In contrast with existing studies, this paper takes a broader view of social influence and demonstrates that characteristics of the community—unrelated to known factors of participation—play a significant role in the financial outcomes of households. Additionally, this paper highlights the importance of soft information for assisting households in overcoming the subjective barriers to participation.

By choosing racial diversity as the primary characteristic of importance, I contribute to the strand of literature working to understand the lack of financial market participation among minorities and the consequences of these low participation rates (Chiteji and Stafford (1999), Blau and Graham (1990), Favilukis (2013), Loury (1998)). These papers propose inter-generational transfers, lack of trust in formal institutions and social learning dynamics as potential drivers of the differences in the asset composition of minority households. This paper demonstrates that effective racial integration among local communities can provide better access to information, effectively lowering the informational costs to participation which encourages minority households to invest in financial markets. Additionally, these findings the cumulative detrimental effects of segregationist policies like residential redlining can help explain the large and persistent gap in wealth accumulation.

Overall, the results presented in this paper provide evidence that the local environment has a significant influence on the financial decisions of households. Recognizing that the external environment has a significant influence on the way individuals obtain information provides a new lens through which to view the household financial decision making process. Furthermore, the finding that integration results in a pareto-improvement in financial outcomes for individuals of all races highlights the significance of social structures in the financial well-being of modern households.

The remainder of this paper is organized as follows. In section II., I discuss the data sources and the derivation of the main variables of interest. Section III. elaborates on the causes and characteristics of racial diversity across the US. Section IV. develops the empirical methodology and presents the main results. Section V. examines whether local integration has differential effects across race. Section VI. examines the robustness of the main findings. Section VII. explores the potential economic channels that connect diversity to the participation decision. Section VIII. concludes.

II. Data Sources and Main Variables of Interest

A. Racial Residential Integration

In order to test the implications of racial integration for household decision making, I must first characterize what is meant by integration. In this context, a measure of integration must be able to capture the *spatial distribution* of individuals within a particular area. More specifically, an integration statistic should provide a single number that characterizes the two dimensional distribution of a population's subgroups across units (White (1986)). Units are defined spatially and can be as large as counties or as small as individual neighborhoods. In the context of this paper, I characterize integrated environments as those in which the likelihood of interracial interaction is high.

Since this paper is concerned with understanding the effects of diverse social networks and interracial interactions, an index of exposure is the best fit to capture the degree of integration within a locale.[3]

[3]Massey and Denton (1989) argue that this measure of exposure makes the most sense in a sociological context because it is better able to describe social experiences of group members in different populations. Additionally, Echenique and Fryer Jr (2007) find support that this measure of segregation is the most closely related to a measure that identifies actual social interaction among friendship networks.

The probability of one group interacting with another is defined as

$$P(exposure) = \sum_{i=1}^{k} (x_i/X)(y_i/n_i) \tag{1}$$

where x_i represents the number of Black individuals in a tract, X represents the total number of Black individuals in the geographic unit, y_i represents the number of White individuals in a tract, and n_i represents the total number of individuals in a tract. These ratios are then summed over all tracts within a county to arrive at the probability that a Black individual interacts with a White individual. I restrict the sample to include only counties that have non-zero Black and White populations.

Panel A of table 1 presents summary statistics on the exposure measure as well as the racial composition of US census counties. Evidence of residential segregation can be seen in the averages of these two statistics. Given that Black Americans represent approximately 13% of the US population, the unconditional probability that Black interacts with White should be close to 87% if the population were evenly distributed. The conditional probability based on the spatial organization of racial groups within neighborhoods, however is only 78%. This discrepancy illustrates the propensity for individuals to form clusters among their own racial groups.

An issue with this measure of interracial exposure, however, is that it is asymmetric with respect to the race of the observer. As it is computed above, the measure of exposure provides the likelihood that Black interacts with White. Because the analysis is based on identifying experienced diversity for individuals of all races, this asymmetry prevents the use of the linear continuous measure of exposure probabilities as a proxy for local diversity. In order to identify a more symmetric measure, the classification of diversity must be such that

individuals from either race should face a similar likelihood of interacting with someone with a different racial background.

Given that I characterize diverse environments as those in which the likelihood of interracial interaction is high—regardless of an individual's race—I define counties with 50% exposure probability as displaying the maximum amount of diversity. In counties where the exposure probability is 50%, individuals of either race are equally likely to interact with those of the same race and those of a different race.

For ease of interpretation, I create a binary classification of integration as follows:[4]

$$Integrated_{jt} = \begin{cases} 1 & \text{if } P_{bwj} = 50\% \pm \frac{1}{2}\sigma \\ 0 & Otherwise \end{cases} \tag{2}$$

Where D_{jt} is a binary indicator for integration at the county-decade level, P_{bwj} is the probability of exposure as in equation 1, and σ is the standard deviation of the probability of exposure. Data on population estimates come from four waves of the US Census from 1980 to 2010. The census provides the total population by race residing in a census tract, which can then be used to calculate the statistic in equation 1, which is subsequently mapped to my empirical definition of integration in equation (2).

Panel B of table 1 presents county level characteristics broken down by integrated and segregated counties. Due to the persistent nature of segregation across the US, the number of segregated counties outnumber integrated counties. Given the skewed nature of the county exposure distribution, a potential concern is that those counties falling within this

[4]All tests are also performed using a continuous measure which is derived by taking the absolute value of the deviation from 50%. Tests are also performed using alternative cutoffs for the binary integration indicators where the band is either narrower or wider than what is presented in eq 2. For specifications run on wider (narrower) bands, the statistical significance is stronger (weaker) while the economic significance is weaker (stronger) indicating a trade-off between power and efficiency in the estimation.

band (around the 40^{th} and 60^{th} percentiles) have special characteristics that may bias the results. Panel B of table 1 presents summary statistics by county classification as well as the differences between each county type.

Lastly, it is important to make the distinction between the unconditional probability of interaction based merely on population statistics and the conditional probability of interaction based on equation 1. The exposure based probability of interaction takes into account the clustering of individuals in census tracts within the county. Therefore, it is possible to have two counties that have the same population distribution (percentage of Black or White in the county) but have different exposure probabilities due to spatial clustering.

An example of this type of relation is the comparison between Queens County, NY and Bedford City, VA. The Black population comprises around 20% of the total county population in both counties, while the probability of Black meeting White is 15% in Queens County and 76% in Bedford City. Such a comparison emphasizes the novelty of the exposure measure in identifying interracial exposure based on the spatial distribution of racial groups within a county.

B. Household Survey Data

To test whether local integration influences household financial decision making, I merge the county-level integration measure derived from the US Census into household-level panel data from the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal household survey conducted by the University of Michigan started in 1968. The public version of the PSID data contains geographic identifiers at the state level. However, because this study requires an understanding of an individual's *local* environment, I acquired access to the

restricted geocoded files containing household locations at the census tract level.[5] Another benefit of this data set is that it also tracks household location across and between panels.

In addition to the granular data on household location, the PSID provides information on a wide variety of topics such as individual attitudes and behaviors, income, wealth, and a host of demographics characteristics. The variable of interest in this study is equity participation. Equity participation is an indicator variable that takes the value of one if the household reports owning stock in a publicly held corporations, mutual funds or investment trusts in a given survey year. This variable is available every five years starting in 1984 and then every other year from 1999 to the most recent survey in 2017. For waves prior to 1999, the question asks households to include stocks held in individual retirement accounts (IRAs). Because investments in IRAs are often associated with default choices, I restrict equity participation to only include non-IRA assets (Choi, Laibson, Madrian, and Metrick (2002)).

Table 2 reports summary statistics for PSID households. All statistics are weighted using the appropriate population weights throughout the analysis. Column (1) presents the statistics for the full sample of households, column (2) ((3)) restricts the sample to include households residing in integrated (segregated) counties in a given survey year. These statistics illustrate that there does not seem to be any economically large systematic differences between households residing in integrated or segregated counties. Differences in observable characteristics are also controlled for in all regression specifications.

[5]Some of the data used in this analysis are derived from Restricted Data Files of the Panel Study of Income Dynamics, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the authors. Persons interested in obtaining PSID Restricted Data Files should contact PSIDHelp@umich.edu.

C. Brokerage Data

To test the aforementioned conjectures, I require detailed data on the financial investments and decisions of households. Given that detailed financial information is limited in the PSID, I utilize individual holdings and trades data from a large discount brokerage firm which reports the investments of 78,000 households from across the United States (refer toBarber and Odean (2000) for a more detailed description of this data set). This data provides the positions held by each household at a monthly level, as well as the quantity of each security traded for households from 1991 to 1996. Therefore, analysis done on individual investor behavior is restricted to this time period.

Since the analysis is primarily concerned with investors' location of residence, I restrict the data set to include only households for which zip code and demographic information is available. I then merge the county-level integration measure into the brokerage data set to identify whether each investor resides in an integrated or segregated community. I also exclude households with residence outside of the continental United States and who have less than $1,000 in their portfolio as of December 1991. This results in a sample of 39,521 unique households. On average, households in the sample hold 3.7 stocks in their portfolios and have a total portfolio value of $44,916. Table 3 presents summary statistics for households in the sample by diversity categorization. In comparing the statistics between panels B and C, it can be seen that the households residing in diverse areas are not qualitatively different from those in segregated areas.

This paper focuses on common stocks traded on the NYSE, AMEX and Nasdaq exchanges. Data on prices and returns come from the Center for Research and Security Prices (CRSP). I obtain zip code level location information and other descriptive characteristics from COMPUSTAT. As in Ivković and Weisbenner (2005) I use firm headquarter location

as opposed to incorporation because firms typically do not conduct the majority of their operations in their state of incorporation. Data on returns for the value-weighted market portfolio, risk-free rate and the Fama-French factors are downloaded from Kenneth French's website.

To measure the concentration in local firms as well as to create locally based calendar-time portfolios, I translate household and firm zip codes into latitudes and longitudes using the 1990 U.S. Census Gazetteer place and zip code database. I compute the arc length distance between households and firms using the formula used by Coval and Moskowitz (2001). I define an investment to be local if the firm is headquartered within 250 miles of the household.

III. Racial Residential Integration

Racial integration is a salient characteristic of the local environment that is likely to influence how individuals think and behave. While integration can be defined along multiple dimensions, this paper isolates the role of Black-White integration in the financial decision-making process. This section provides a brief historical background as context for understanding the basis of modern social structures.

A. Historical Context

Following the abolition of slavery in the United States in 1865, Black Americans were geographically concentrated in the southern states. While technically freed, most were forced to continue to work the land for extremely low wages and had little opportunity for upward mobility. Conditions for Black Americans in the south remained poor throughout the end of the 1800s, providing little opportunity for economic advancement.

Upon the advent of Word War I, factory workers in the north were called overseas to fight, and European immigration to the United States subsided. As jobs became available in industrial centers in the north, companies began to actively recruit Black Americans in the south to fill these positions. With the hope of new opportunities, thousands of Black Americans travelled north and west, thus beginning the first wave of the Great Migration. The migration slowed at the end of the first world war, however it picked up again in similar fashion during the course of World War II. Over the period of 1916 to 1970, an estimated 6 million Blacks left the south (US Census Bureau (2012)). Figure 1 illustrates the extent and variation of the migration patterns over the course of the migration.

The Great Migration introduced a dramatic shift in the demographic landscape across the country. Northern cities that experienced large population shocks reacted with discriminatory and segregationist policies that created long lasting racial tensions in some areas. Conversely, areas that experienced less of an extreme shock did not enact such segregationist policies, resulting in a higher degree of racial integration. The heterogeneity in both the location choices as well as the reactions to new migrants led to significant variation in the demographic composition of cities and the racial attitudes in them.

While legislation passed over fifty years ago such as the Fair Housing Act attempted to reduce residential segregation, this phenomena is still prevalent today. Theories regarding the persistence of residential segregation range from homophily and discrimination (Quillian (2002), Charles (2003)) to public policy initiatives (Abramovitz and Smith (2021)). Regardless of the reason for its persistence, residential segregation plays a fundamental role in racial inequality. This paper is the first to establish how such social structures influence financial outcomes in order to identify potential solutions to racial wealth inequality.

IV. Portfolio Allocation & Community Effects

A. Empirical Methodology

I begin by relating the participation decision to local integration using the following empirical specification:

$$Participate_{ijt} = \beta \times Integrated_{jt} + \gamma X_{it} + \delta Y_{jt} + \zeta_{jt} + \varepsilon_{ijt}, \tag{3}$$

Where *Participate* takes the value of one for households that indicate that they own publicly traded stocks. Integrated is an indicator that takes the value of one if the county is classified as integrated as defined in equation (2), X includes a vector of time-varying household level controls, Y contains a vector of time-varying county level controls and ζ represents state-by-year fixed effects. Household level controls include: (i) indicator variable for each percentile of the wealth distribution by year (the 50th percentile is the omitted category) (ii) an indicator variable of gender interacted with filing status (single male, single female, married male, married female) as well as the stand alone variables (iii) indicator variables for the degree of risk aversion and connectedness (iv) other demographic items such as race, age, age^2, education, income and income risk.[6]

An important issue that must be addressed in running this specification is identifying characteristics of the local environment that may be correlated with both racial integration as well as the participation decision. Apart from the typical individual level controls known to be associated with the participation decision (education, income, wealth, etc.), this model also includes county level characteristics that account for local socioeconomic conditions.

[6]The PSID contains questions to characterize household attitudes and behaviors. The question on connectedness contains an index based on the degree of connectedness to help. This index contains items such as whether the individual attends PTA meetings, attends church, watches the news, knows their neighbors and attend social organizations.

County-level socioeconomic controls include: (i) an indicator for whether the county is classified as urban (ii) natural log of median home value (iii) percentiles of county population (iv) a continuous measure of social capital from Rupasingha, Goetz, and Freshwater (2006) (v) fraction of individuals in county j that participate in the stock market excluding household i (vi) fraction of individuals working in the finance industry (vii) averages of individual level controls (race, age, education, etc.). In adding these controls, I hope to belay the concern that the coefficient on *Integrated* is biased by some other community-level characteristic unrelated to the presence of racial integration.

Additionally, following Brown, Ivković, Smith, and Weisbenner (2008), I include a vector of controls that characterize the conditions among locally publicly traded firms. Local financial market controls include (i) the presence of publicly traded firms in the county (ii) the fraction of total US market value represented by firms in county j (iii) value weighted return of firms in county j and (iv) the fraction of firms in the county paying out dividends. These controls are intended to account for any variation caused by local financial markets that may contribute to the likelihood that a household residing in the county would participate in the stock market. Finally, all specifications include state-by-year fixed effects to control for any systematic differences across states and across time. The variable definitions table in the appendix contains more detail regarding the explicit calculation of each control variable as well as the data source from which it was obtained.

Also following Brown, Ivković, Smith, and Weisbenner (2008), I estimate this model using ordinary least squares.[7] Standard errors are clustered at the county level to allow for correlation across individuals within each county as well as across individuals in the same county over time. Given the large number of control variables, there is a concern that

[7] As robustness, column (1) of appendix table A2 presents results from the model estimated using a probit specification which yields similar results.

multicollinearity may inflate the variance estimates of the model. A check of the correlations as well as variance inflation factors (VIF) across all controls alleviates the concern that there is severe multicollinearity driving the significance of the results.

B. Baseline Results

Table 4 presents the baseline results. The coefficient on the *Integrated* indicator is positive and statistically significant, indicating that a household is significantly more likely to participate in the stock market if they reside in a community where interracial interactions are common. The coefficient on *Integrated* in column (1) is 0.01 and with a significance level of 95%. Given the unconditional participation rate of 24%, the estimate suggests that integrated investors are about 4% more likely to invest in publicly traded equities. The magnitudes and significance of the coefficients remain relatively unchanged whether I include individual level and county level sociability measures (connectedness and community ownership respectively), which indicates that integration is not subsumed by factors of sociability found in prior studies.

The economic magnitude is comparable to recent studies examing the stock market participation decision. For example, Brown, Ivković, Smith, and Weisbenner (2008) find that a 10% increase in in community ownership is associated with a 4.2% increase in participation likelihood. Similarly, Giannetti and Wang (2016) find that a one-standard-deviation increase in lifetime exposure to fraud reduces the probability of a household owning stock by about 4%.

The effect of diversity measured here can also be seen as a lower bound for the effects of local environments on participation. Measuring local integration in terms of other races or in terms of other characteristics such as gender and income may strengthen these effects.

C. Identification from Household Relocation

While the extensive list of controls helps to eliminate unwanted variation due to characteristics of the individual and the local environment, there remains a concern that there may be some unobservable factors contaminating the results. The existence of an unobservable omitted variable that is correlated with both the degree of racial integration in the county as well as the equity participation decision would result in a bias in the baseline results. In this section, I outline the first of two identification strategies aimed at addressing this concern.

The first strategy utilized to alleviate endogeneity concerns is to identify within individual variation by testing whether households change their behavior after moving from a segregated to an integrated county. For this set of tests, I restrict the sample to include households that either (i) reside in segregated counties throughout the entire sample or (ii) begin the sample residing in segregated counties and move once to an integrated county and remain in this county throughout the remainder of the sample. For the first condition, I include households that may move from one segregated to another segregated county because restricting to only segregated non-movers may elicit sample selection since I will be comparing these households to "mover" households. Restricting the sample in this way provides for a cleaner comparison between households who tend to maintain residence in segregated counties to those who moved to integrated counties. By restricting movers to those who began the sample in a segregated county, this alleviates some of the concern that certain types of people self select into either segregated or diverse communities which allows for more of an apples to apples comparison.

Because I am identifying the effects of moving to integration, I run the tests using the change in participation and include first-differences in all control variables to ensure that results are not driven by changes in other observable characteristics such as income or other

life events (i.e. job promotion, marriage, etc.). More specifically I test the effect of a move to integration at time t on the change in the participation decision between $t-1$ and t. Destination-level socioeconomic and financial market controls are used in this specification. I also explore the persistence of the effects of moving to an integrated community over time. To do so, I test whether the participation decision is affected up to six years following the move.

Table 5 presents the results from the moving to integration specification. The positive and significant coefficient on the moving to integration indicator in column (1) demonstrates that individuals are 1.2% more likely to participate following a move to an integrated county. Column (2) looks at the participation decision in the survey year following the move to integration, meaning the move occurred at time t and I am analyzing the change in participation between time $t-1$ and $t+1$.

The slightly larger coefficient on moving to integration in column (2) indicates that individuals are even more likely to participate in the period following the initial move, implying that the effects of living in an integrated county strengthen with time. Intuitively, it takes time for households to acclimate to their surroundings and to develop friendship networks. While the statistical significance is reduced, the economic significance of the effect of the move to integration is even larger in column (3) which further supports the notion that the benefits of soft information present in racially diverse counties compound over time. The effect of integration loses significance six years following the move. This loss in significance is likely partially a result of losing statistical power due to the relatively short time period available in the data.

These findings provide evidence that racial integration has a significant influence on behavior and these effects persist up to four years after the move. This highlights the

ability of the experienced environment to affect the decisions and outcomes of households. Overall, the evidence provided in this section suggests that the experienced environment is a significant contributor to the participation decision. Exploiting the panel structure of this data set allows me to identify within individual variation that mitigates the chance that results are driven by endogenous selection of a certain type of household into racially integrated counties.

D. Identification from Historical Population Shocks

Even after controlling for an extensive set of controls and identifying within individual variation from moves, there is still a possibility that the individuals in the sample are sorting themselves into certain environments based on some unobserved factor that also happens to be correlated with the choice to participate. In a best effort to isolate variation due to the component of interest in this study—local racial integration—I implement an instrumental variables approach to reduce additional endogeneity concerns.

More specifically I instrument for contemporaneous local integration using the predicted Black population shock to northern commuting zones (CZs) during the Great Migration between 1940 and 1970 (Derenoncourt (2019)). A proper instrument for this study would be one which identifies exogenous variation in the degree of contemporaneous local diversity while being uncorrelated with the outcome of interest. The benefit of using a *predicted* population shock is that it cleanly identifies population shocks by reducing the influence of omitted variables that might have influenced socioeconomic outcomes in destination areas. In other words, this instrument disentangles exogenous variation in population shocks from the endogenous selection of families into northern commuting zones.[8]

[8]Boustan (2010), Black, Sanders, Taylor, and Taylor (2015), Stuart and Taylor (2019) demonstrate that there exists specific linkages between southern and northern location choices during the Great Migration.

The data used to construct the predicted Black population shock is borrowed from Derenoncourt (2019), but is originally drawn from the 1940 complete count US census and the City and County Data Books for the years 1944-1977. The sample is restricted to destination commuting zones (northern and western regions) for which there is available data. This results in a final sample of 130 commuting zones with one observation per CZ to represent the population shock between 1940 and 1970.

The instrument is created by interacting destination northern cities' pre-1940 migrant composition with variation in outmigration from southern counties driven by only push factors (factors that pushed Black migrants out of southern CZs). More explicitly, the *predicted* Black population shock is computed as

$$Predicted\, black\, pop_{CZ}^{1940-1970} = \frac{\hat{\Delta} b_{CZ}^{1940-1970}}{pop_{CZ}^{1940}} \tag{4}$$

where $\hat{\Delta} b_{CZ}^{1940-1970}$ is the predicted increase in the Black population for each CZ during the period 1940-1970. This variable is predicted by interacting the pre-1940 migrant composition in northern CZs with predicted southern county net migration over the entire Great Migration period (1940-1970). Predicted southern net migration is computed using a Post-LASSO estimation procedure to isolate the most important predictors for southern net migration.[9] This predicted population shock is then transformed into percentiles given the severe right skew of these population shocks.

Before moving on to the empirical implementation of this approach, it is important to note how these population shocks impacted northern and western communities in order to understand how it relates to contemporaneous racial integration. In some urban destinations, the large Black population shocks engendered discriminatory reactionary policies that

Such linkages may result in endogeneity with respect to migration and economic outcomes in destination locations.

[9] Please refer to Derenoncourt (2019) for a more detailed explanation of the construction of this instrument.

resulted in residential segregation and rising racial tensions. Alternatively, some areas did not attract large population flows, leaving the demographic mix unchanged. Therefore, I conjecture that CZs that experience moderate population shocks are representative of areas characterized by high interracial interactions and can serve as a good instrument for contemporaneous integration.

While these areas experienced a moderately large inflow of Black residents during this time, the extent of the shocks in these areas did not elicit the backlash that was seen in the CZs that experienced larger shocks, which allowed for more effective integration in these destinations. Following this conjecture and mirroring the methodology defining contemporaneous racial integration, I classify CZs as being integrated if the fall within the 40-60th percentile of the predicted population shock and segregated if they are outside of this band. For the remainder of the discussion I classify CZs that fall within the 40-60th percentile as moderate shock CZs and those that fall outside of this cutoff extreme shock CZs.

While the relevance condition is intuitive for this instrument and can be tested in the first-stage regression, the exclusion restriction requires a more in depth discussion. I argue that the exclusion restriction for this instrument holds because it is unlikely that historical population shocks meaningfully impact an individual's contemporaneous decision to participate. Panel A of table 6 illustrates that the differences in observable characteristics are not economically large across moderate and extreme shock CZs. Extreme shock (segregated) CZs actually have higher median household incomes and home values which would result in a bias the results in the other direction. After accounting for observable CZ and individual level characteristics, there does not seem to be obvious violations of the exclusion restriction.

Panel B of table 6 presents the results from the first stage regression. The strong significance on the coefficient in column (1) illustrates that moderate predicted Black population

shocks are strongly correlated with contemporaneous racial integration. The validity of the instrument is also supported by the high first stage F-statistic. The second stage estimates are presented in column (2). The exogenous variation in racial integration is shown to significantly increase the likelihood of stock market participation. The economic significance of the second stage results further supports the claim that experienced racial integration at the local level impacts a household's decision to participate in the stock market. By isolating only the exogenous variation due to contemporaneous integration, this specification is better able to tease out the effects of integration on participation.

V. Differential Effects Among Demographic Subgroups

Given that integration is found to increase participation among the general population, this section explores whether these effects differ across demographic groups. More specifically, this section asks whether the integration effect benefits one race over another, or whether all individuals can reap the benefits. The motivating issue for this paper is to understand factors that prevent minorities from participating or factors that may induce minorities to participate. In this context, it is important to understand exactly who is benefiting from these social structures.

A potential explanation for the positive impact of integration on participation is, given that White Americans are statistically more likely to invest, that information could flow from the participating majority to the non-participating minority. If this were the case, the coefficient on *Integrated* should be significant for the Black population and not for White. Alternatively, if community integration facilitates information transmission or contributes to a change in preferences in a way that is not specific to the race of the individual, then the marginal effects of integration should be significant for both subgroups.

To identify the marginal effects of integration across racial groups, I introduce an interaction of race with integration in the baseline specification. Specifically, I interact the county-level integration indicator with an indicator taking the value of one if the survey respondent identifies as Black. A positive (negative) coefficient on this interaction term will show that integration affects the participation rate of Black individuals significantly more (less) than White individuals. An insignificant coefficient demonstrates that there is no significant difference in how diversity affects either race.

I also introduce the interaction term in the moves specification from section III..C.. I interact the race indicator with each of the moving to integration indicators. This specification tests whether the effects of moving to integration differ across demographic groups. The interpretation of the interaction term is similar to the baseline specification in that it demonstrates whether the effects of moving to integration differentially affects White or Black individuals.

Figure 2 illustrates the results. Panel (a) plots the predicted participation rates by integration and race for the baseline interaction specification. A comparison across integrated and segregated categories demonstrate the baseline result that integration is associated with higher participation rates. Important to note is that integration significantly increases participation rates for both White and Black individuals. This demonstrates that integration results in a pareto-improvement in financial outcomes and that integration benefits all individuals.

While the predicted participation rates are higher for both Black and White individuals, the marginal effect of integration is not significantly different across subgroups. The difference in participation rates for individuals residing in integrated communities is not larger

for one group relative to the other. In other words, both groups seem to be able to extract the same amount of beneficial information from integrated environments.

As with the baseline specification presented in panel (a) of figure 2, I plot the predicted participation rates after the moving to integration specification with the added an interaction term in panel (b). As was the case with the baseline specification, the interaction coefficient does not load significantly indicating that there is no significant difference in the effect of moving to an integrated county between White and Black households. However, the difference in participation rates between those that move to diversity and those that stay in racially isolated communities is positive and significant for both White and Black households.

Overall, the interesting takeaway from the results above is that integrated communities not only improve financial outcomes for minorities, but it also increases the likelihood of participation for majorities. Recognizing local integration as having a systematic influence across all households emphasizes the importance of such local factors.

VI. Robustness

In this section, I present additional evidence to solidify the claim that community-level racial integration is consequential for the participation decision.

A. Integration Effects in the Cross-Section

To provide additional evidence that racial integration is associated with individual equity ownership, I test whether the effect of integration is stronger in areas where people of different races are more likely to interact with each other one-on-one. While the integration measure picks up the likelihood of interaction based on racial clustering among census tracts, the

degree of sociability in an area is likely to increase the strength of the integration effect if relevant information is gained through interpersonal interactions.

I select three community-level characteristics that are associated with increased personal interaction among community members. The first measure is the county's population density. Intuitively, the denser an area, the more likely individuals are to come into contact with one another on a recurring basis. The data on county-level population density comes from the U.S. Decennial Census and is computed as the number of residents divided by the total land area.

The second and third measures are the number of civic organizations and recreational sports centers in a county. This data comes from Rupasingha, Goetz, and Freshwater (2006). The sociability indices from this data set are a continuation of the social capital data from the DDB Life Style data used in Putnam et al. (2000). This data set provides the total number of each type of organization in a county over time. While this data set provides counts of other types of organizations (such as religious and political groups), civic organizations and recreational centers are more likely to foster interactions across racial groups. The average county has around eight (ten) civic (fitness) organizations.

Table 7 presents the results of specifications that interact the integration indicator with each of the above characteristics. The coefficients on all three interactions are positive and significant. This finding indicates that the marginal effect of diversity is stronger in communities that are better structured for social interaction. These findings support the word-of-mouth effects found in Brown, Ivković, Smith, and Weisbenner (2008) while also emphasizing the fact that the context of the social interaction matters. Effectively, the information gained from social interactions is more influential when the interaction is between individuals from different racial backgrounds.

B. Alternative Definitions of Integration

The main hypothesis of this paper relies on the interaction of racially dissimilar peers. A potential concern is that the integration measure used in the above specifications does not properly identify the likelihood of a person interacting with another of a different race. The exposure index, from which the integration measure is derived, is based on the residential locations of each race across census tracts. While this measure of exposure is a good proxy for racial integration, I augment the analysis by using a novel measure of segregation based on the social networks of individuals.

I utilize the spectral segregation index (SSI) from Echenique and Fryer Jr (2007), which takes into account the segregation of one's social network. The computation of this segregation index makes it such that an individual is more segregated the more segregated her network is. By measuring the degree of racial segregation at the individual network level before aggregating it up to the MSA level, this index ensures that the level of segregation accurately represents the social interactions of individuals in the community.

The data used to compute this index comes from National Longitudinal Study of Adolescent Health (Addhealth) database.[10] The SSI is computed for each individual and each race and is then aggregated to the MSA level. The MSA level SSI therefore provides the degree to which individuals are isolated within their own race.

Because SSI is computed in a way to ensure monotonicity in segregation across all races (as opposed to only Black and White in the exposure measure used to compute integration), I am able to substitute the integration indicator in the baseline specifications for the continuous SSI measure. This index can also be computed in terms of race-specific isolation. For instance, a Black (White) spectral segregation index provides the degree to which Black

[10]The index is available for download on Federico Echenique's website here.

individuals associate only with other Black (White) individuals. I utilize the overall SSI as well as the Black and White SSI to demonstrate that interracial interactions significantly affect financial decisions of households.

An important distinction to note in this substituted model is that a higher value of SSI indicates a higher degree of segregation (i.e. less integration). Therefore, I should expect the sign of the coefficients on SSI to be reversed relative to the integration indicator.

Table 8 confirms this expectation. The coefficient on SSI is negative and significant for the baseline specification as well as the moving to diversity specifications. This demonstrates that participation rates are significantly lower in areas where individuals are more racially isolated. Moreover, the economic magnitude implies a one standard deviation increase in racial isolation results in around a 3% reduction in participation. Compared to the 1% difference in participation rates found in the baseline model, this network-based measure of segregation based on social interactions demonstrates that the participation decision is more sensitive to a measure of racial isolation that takes into account primary and secondary friendship networks.

VII. How Does Integration Influence Participation?

Until now, this paper has worked to establish that exposure to interracial interactions at the community level has a significant impact on household financial decision making. In this section, I explore the channel through which these effects occur.

The main conjecture of this paper is that racially integrated environments improve information environments by increasing diversity among social networks. Integrated environments are likely to foster more diverse social networks which has been found to increase the

likelihood of information transfer across networks (Eagle, Macy, and Claxton (2010)). This effect results from an increased likelihood that individuals have social ties that bridge across groups. In facilitating these social bridges, information that is available in one group in an integrated environment is more likely to spread to other groups compared to a segregated environment where groups are isolated from each other.

I investigate whether integration influences participation via the information channel by testing whether investors exposed to integrated environments achieve superior risk-adjusted portfolio performance relative to those exposed to segregated environments. If value-relevant information is more available in integrated environments, investors should be able to capitalize on this informational advantage and outperform investors with more isolated information sets.

I then explore alternative explanations for the relation between integration and participation. A potential explanation for the positive effect of integration on participation is that integrated areas are associated with greater financial literacy. While the individual and county-level controls attempt to capture characteristics associated with financial literacy, higher participation rates in combination with superior performance suggests that integration could be a proxy for sophistication. To rule out financial sophistication as an alternative explanation, I compare the trading behavior of investors across integrated and segregated environments.

A. Does Integration Provide an Informational Advantage?

Given that integrated investors exhibit a stronger degree of confidence in their investment choices and an increased propensity to take on risk, I test whether this behavior is rewarded with superior performance. If investors residing in integrated communities are more likely to

participate because they have better access to information, I would expect them to achieve higher risk-adjusted portfolio returns relative to racially isolated investors.

Using standard portfolio performance analysis, I first test whether there is a difference in overall performance. I then go on to determine whether there is a difference in performance of the local portion of investors' portfolios. There is an existing debate in the literature regarding whether local investors can actually profit off of local information (Seasholes and Zhu (2010), Ivković and Weisbenner (2005)). In this set of tests, I attempt to uncover whether there are certain conditions under which local information can be successfully extracted for profit. Additionally, I want to understand whether the risky behavior found in the previous section can be explained by these investors having superior information or if it is due to diversity increasing individuals' risk appetites.

For each individual, I first calculate the value-weighted portfolio return for each month. I then create holdings-based calendar-time portfolios of local stocks. To compute local portfolio returns, I compute the value-weighted monthly return for holdings located within a 250-mile radius from the household.[11] Because there might be a concern that certain regions may have experienced abnormally high or low returns over the course of the 1991-1996 sample period, I create a "local portfolio" benchmark by calculating the value-weighted return for all firms located within 250 miles of each zip code.

I then regress the excess portfolio returns on the integration indicator as well as progressive specifications controlling for market risk, the Fama-French factors and the local benchmark. The coefficient on the integration indicator provides the 'integration alpha'.

[11] Alternative definitions of local are also used. Local is alternatively defined by a 100-mile and 50-mile radius obtaining similar results

1. Portfolio Performance

To test whether integrated investors achieve superior risk-adjusted returns, I regress each individual's value-weighted returns (both overall and local) on the indicator for integration as well as several risk-adjusting factors. All regressions control for individual characteristics as well as state-by-year fixed effects to control for any state-level time trends in economic conditions that may bias the results. Table 9 presents the results.

Panel A of table 9 presents summary statistics on household portfolio returns and Sharpe ratios. The first two columns provide the statistics for the overall portfolio and the last two columns provide the statistics for the local portion of the household portfolio. The overall and local excess returns are significantly higher among integrated households. However, only the local Sharpe ratio is significantly higher for integrated investors providing the first indication that these investors may have superior information about local firms.

Panel B of table 9 provides the results on overall performance. The coefficient on the integration indicator is insignificant across all models, indicating that there is no significant difference in the performance of investors residing in integrated versus segregated zip codes.

However, turning to panel C of table 9 I find that the local portfolios of integrated investors significantly outperform their racially isolated counterparts. The coefficients on the integration indicator are significant across all models demonstrating that, controlling for observable characteristics, the *local* portfolios of investors exposed to racial diversity significantly outperform those residing in more racially segregated counties. The evidence supports the notion that integrated investors are able to extract valuable information about local firms, but not necessarily about non-local firms. If it were the case that integrated investors were more sophisticated or simply better investors, I would expect there to be a significant difference in the overall performance as well.

These findings suggest that local integration improves the quality of the local information environment, lowering informational barriers to participation. Ultimately, evidence from trading behavior and performance demonstrate that racially integrated communities enable more effective information sharing, which allows individuals to access value-relevant information that improves financial outcomes.

2. Patterns in Trading Behavior

Superior local portfolio performance among integrated investors suggests that racial integration facilitates higher quality information environments, allowing individuals to make more informed decisions. I next examine whether exposure to racial integration affects the propensity to display irrational or biased tendencies in trading to disentangle whether superior performance is due to investors having information specifically relevant to local firms or if these environments are associated with less biased trading behavior.

I first explore whether the degree of diversification differs across groups. I construct several measures to describe the extent to which investors hold well-diversified portfolios. The first and simplest measure of diversification is the total number of stocks held in each individual's portfolio. I compute this statistic at the individual-month level. On average, investors in diverse zip codes hold 3.6 stocks compared to 3.8 for those residing in segregated zip codes. I next compute the normalized variance of each household's portfolio over the entire sample period to get an idea of the level of sophisticated diversification. Normalized variance is obtained by dividing the portfolio variance by the average variance of all stocks in the portfolio. The covariance matrix is computed using the past five years of monthly returns.

I then compute two separate measures to understand whether diversity exposed investors trade more aggressively or have a greater tendency to display one of the most widely cited

behavioral biases—the disposition effect. The first is average portfolio turnover. Investors that turnover their portfolios more frequently likely believe that their bets will be lucrative enough to compensate for the additional trading costs. This confidence in trading ability is also associated with an increased willingness to take risk (Merkle (2017), Menkhoff, Schmidt, and Brozynski (2006)). Portfolio turnover is the average of monthly buy and sell turnover. This statistic is computed as detailed in Barber and Odean (2001).

I then compute the disposition effect as an additional measure to evaluate the degree of biased tendencies. Following the data definitions of Ivković, Poterba, and Weisbenner (2005), I compute the disposition effect as the difference between the propensity to realize gains (PGR) and the propensity to realize losses (PLR).[12]

Finally, I determine the degree of local bias for each investor. I compute the fraction of each investor's portfolio that is local (within 250 miles of the household) to get a sense of relative concentration in local firms. To account for the fact that households residing near financial centers are more likely to have a higher concentration of local stocks, I compute a benchmark for the local portion of a portfolio by computing the fraction of total market value within 250 miles of each zip code. A local bias is present if a household holds a significantly larger fraction of local stocks in their portfolio relative to what would be held in an equally weighted market portfolio.

To empirically test whether integrated and segregated investors display differences in trading behavior, I regress each of the above measures on the integration indicator as well as a vector of individual-level controls available in the brokerage data set.

[12]Where PGR (PLR) is calculated as the number of realized gains (losses) divided by the sum of the number of realized and paper gains (losses).

Table 10 presents the results. These results indicate that investors residing in integrated counties are relatively less diversified and trade more aggressively than racially isolated investors. Integrated investors hold significantly fewer stocks in their portfolios, have higher normalized variance and higher average turnover. Additionally, in untabulated results, I find that integrated investors also have higher raw portfolio standard deviations indicating a higher tolerance for risk.

Additionally, adjusting for the fraction of the market that is headquartered locally, column (6) of table 10 illustrates that integrated investors concentrate a larger fraction of their portfolio in local stocks. Together, these patterns in trading behavior suggest that integrated investors are surprisingly more likely to exhibit behavioral biases along some dimensions. Greater concentration in local stocks in addition to lower diversification and higher turnover suggest that they are more confident in their trading abilities. While these behaviors are typically associated with irrationality and poor performance, this finding in combination with superior local performance provides support for the conjectured information channel.

VIII. Conclusion

This paper demonstrates that racial integration has significant implications for investor behavior through its effect on the information environment. Connecting residential segregation to investor behavior allows for a better understanding of how social structures interact with financial markets. By utilizing a measure of integration that isolates the probability of interracial interaction from the unconditional probability based merely on population estimates, I am able to provide unique insight into how everyday exposure to integration can influence financial decision making.

The baseline findings are supported by a specification employing within-individual variation testing whether moving to an integrated county increases the likelihood of participation. The main findings are also confirmed by an instrumental variables specification that instruments contemporaneous integration with the predicted Black population shocks during the Great Migration of 1940-1970. While each test separately does not fully control for all endogeneity concerns, the specifications taken together provide convincing evidence that racial integration significantly improves the likelihood of participation at the local level.

Evidence presented on the impact of residential integration on individual investor behavior indicates that the local environment that an individual is exposed to can significantly affect financial outcomes. The notion that an individual can be more or less likely to access returns provided by financial markets depending on where they live is a novel insight that adds a new dimension to the stock market participation puzzle. The fact that integration has an equally positive impact for both Black and White households also emphasizes the systematic importance of these local effects.

Furthermore, I demonstrate that the relation between local racial integration and the participation decision is mediated by an improvement in the quality of local information. The finding that integrated investors are able to achieve superior risk-adjusted performance on their local portfolios suggests that they are able to capitalize on the increased availability of information that is relevant to local firms. Overall, the findings in this paper highlight the importance of local social structures for financial decision-making and outcomes and help to lay groundwork for future research.

References

Abramovitz, M., and R. J. Smith, 2021, "The Persistence of Residential Segregation by Race, 1940 to 2010: The role of federal housing policy," *Families in Society*, 102(1), 5–32.

Addoum, J. M., S. Delikouras, D. Ke, and G. M. Korniotis, 2018, "Clustering Fosters Investment: Local Agglomeration and Household Portfolio Choice," .

Barber, B. M., and T. Odean, 2000, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *The journal of Finance*, 55(2), 773–806.

——— , 2001, "Boys will be boys: Gender, overconfidence, and common stock investment," *The quarterly journal of economics*, 116(1), 261–292.

Behrman, J. R., O. S. Mitchell, C. K. Soo, and D. Bravo, 2012, "How financial literacy affects household wealth accumulation," *American Economic Review*, 102(3), 300–304.

Black, D. A., S. G. Sanders, E. J. Taylor, and L. J. Taylor, 2015, "The impact of the Great Migration on mortality of African Americans: Evidence from the Deep South," *American Economic Review*, 105(2), 477–503.

Blau, F. D., and J. W. Graham, 1990, "Black-white differences in wealth and asset composition," *The Quarterly Journal of Economics*, 105(2), 321–339.

Bonaparte, Y., and A. Kumar, 2013, "Political activism, information costs, and stock market participation," *Journal of Financial Economics*, 107(3), 760–786.

Bonaparte, Y., A. Kumar, and J. K. Page, 2017, "Political climate, optimism, and investment decisions," *Journal of Financial Markets*, 34, 69–94.

Boustan, L. P., 2010, "Was postwar suburbanization white flight? Evidence from the black migration," *The Quarterly Journal of Economics*, 125(1), 417–443.

Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner, 2008, "Neighbors matter: Causal community effects and stock market participation," *The Journal of Finance*, 63(3), 1509–1531.

Campbell, J. Y., 2006, "Household finance," *The journal of finance*, 61(4), 1553–1604.

Charles, C. Z., 2003, "The dynamics of racial residential segregation," *Annual review of sociology*, 29(1), 167–207.

Chiteji, N. S., and F. P. Stafford, 1999, "Portfolio choices of parents and their children as young adults: Asset accumulation by African-American families," *American Economic Review*, 89(2), 377–380.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick, 2002, "Defined contribution pensions: Plan rules, participant choices, and the path of least resistance," *Tax policy and the economy*, 16, 67–113.

Coval, J. D., and T. J. Moskowitz, 2001, "The geography of investment: Informed trading and asset prices," *Journal of political Economy*, 109(4), 811–841.

Derenoncourt, E., 2019, "Can you move to opportunity? Evidence from the Great Migration," *Unpublished working paper*.

Eagle, N., M. Macy, and R. Claxton, 2010, "Network diversity and economic development," *Science*, 328(5981), 1029–1031.

Echenique, F., and R. G. Fryer Jr, 2007, "A measure of segregation based on social interactions," *The Quarterly Journal of Economics*, 122(2), 441–485.

Favilukis, J., 2013, "Inequality, stock market participation, and the equity premium," *Journal of Financial Economics*, 107(3), 740–759.

Giannetti, M., and T. Y. Wang, 2016, "Corporate scandals and household stock market participation," *The Journal of Finance*, 71(6), 2591–2636.

Granovetter, M., and R. Soong, 1983, "Threshold models of diffusion and collective behavior," *Journal of Mathematical sociology*, 9(3), 165–179.

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

Guiso, L., P. Sapienza, and L. Zingales, 2018, "Time varying risk aversion," *Journal of Financial Economics*, 128(3), 403–421.

Hong, H., J. D. Kubik, and J. C. Stein, 2004, "Social interaction and stock-market participation," *The journal of finance*, 59(1), 137–163.

Ivković, Z., J. Poterba, and S. Weisbenner, 2005, "Tax-motivated trading by individual investors," *American Economic Review*, 95(5), 1605–1630.

Ivković, Z., and S. Weisbenner, 2005, "Local does as local is: Information content of the geography of individual investors' common stock investments," *The Journal of Finance*, 60(1), 267–306.

Ke, D., 2021, "Who Wears the Pants? Gender Identity Norms and Intrahousehold Financial Decision-Making," *The Journal of Finance*, 76(3), 1389–1425.

Loury, G. C., 1998, "Why more Blacks dont invest," *New York Times*, 6.

Malmendier, U., and S. Nagel, 2011, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*.," *Quarterly Journal of Economics*, 126(1), 373 – 416.

Massey, D. S., and N. A. Denton, 1989, "Hypersegregation in US metropolitan areas: Black and Hispanic segregation along five dimensions," *Demography*, 26(3), 373–391.

Menkhoff, L., U. Schmidt, and T. Brozynski, 2006, "The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence," *European economic review*, 50(7), 1753–1766.

Merkle, C., 2017, "Financial overconfidence over time: Foresight, hindsight, and insight of investors," *Journal of Banking & Finance*, 84, 68–87.

Putnam, R. D., et al., 2000, *Bowling alone: The collapse and revival of American community*. Simon and schuster.

Quillian, L., 2002, "Why is black–white residential segregation so persistent?: Evidence on three theories from migration data," *Social science research*, 31(2), 197–229.

Rupasingha, A., S. J. Goetz, and D. Freshwater, 2006, "The production of social capital in US counties," *The journal of socio-economics*, 35(1), 83–101.

Seasholes, M. S., and N. Zhu, 2010, "Individual investors and local bias," *The Journal of Finance*, 65(5), 1987–2010.

Stuart, B., and E. J. Taylor, 2019, "Migration networks and location decisions: Evidence from us mass migration," working paper, IZA Discussion Papers.

US Census Bureau, X., 2012, "The Great Migration, 1910 to 1970," .

White, M. J., 1986, "Segregation and diversity measures in population distribution," *Population index*, pp. 198–221.

Figure 1: Population Dynamics of the Great Migration

This figure illustrates the change in share of Black migrants during the two major periods of the Great Migration. This figure is borrowed from the Census and can be found here. The figure on the left graphs the change in the Black population share over the course of the first wave of the Great Migration. The figure on the right graphs the change in the Black population share over the course of the second wave of the Great Migration.

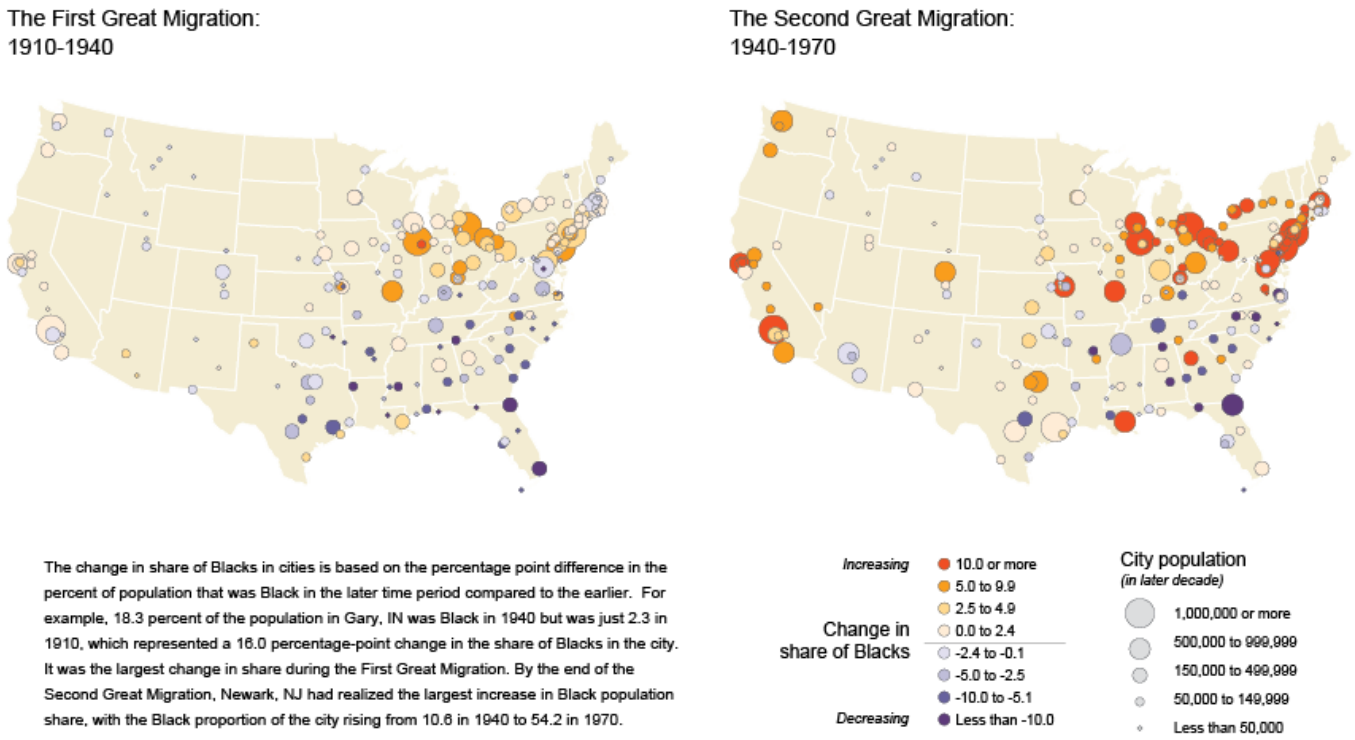
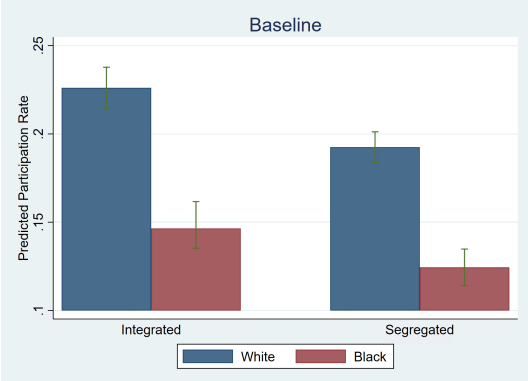
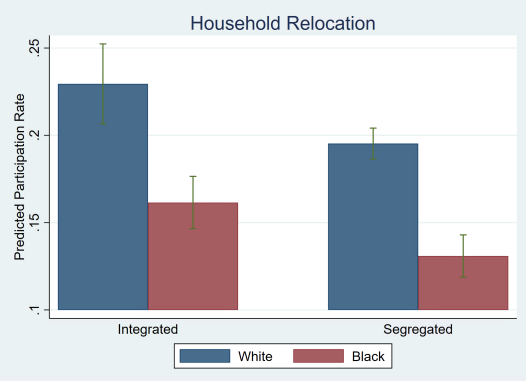


Figure 2: Predicted Participation by County Type and Race

This figure shows the predicted participation rates by race. Panel (a) provides predicted participation rates derived from the baseline specification. The bars represent the predicted participation rates for White and Black households across integrated and segregated counties. Panel (b) presents the predicted participation rates by race after performing the relocation specification where the dependent variable is the participation decision during the survey year of the indicated move. The "Integrated" bars represent the predicted participation rates for households that moved from segregated to integrated counties in the year following relocation. The "Segregated" bars represent the predicted participation rates for households that remained in segregated counties over the same time period. The error bars convey 95% confidence intervals around the predicted values.



(a) Predicted Participation: Baseline Specification



(b) Predicted Participation: Relocation Specification

Table 1: County-Level Statistics For Integration Measure

This table presents county-level summary statistics. Panel A presents statistics on county exposure which is computed as in equation 1. The sample is restricted to counties reporting positive population estimates for both Black and White racial groups. Panel B reports county-level demographic characteristics broken down by integration classification as in equation 2. These variables are obtained from the US Decennial Census and are averaged over the four census waves spanning 1980-2010.

Panel A: County Exposure and Demographic Composition

	Mean	Median	SD	Min	Max	N
County Exposure	78.80%	85.89%	20.36%	2.69%	99.98%	11,131
% Black county	9.35%	2.22%	14.63%	0.01%	94.15%	11,131
% White county	84.68%	90.98%	16.30%	2.69%	99.95%	11,131

Panel B: Demographics by Integrated and Segregated Classification

	Integrated	Segregated	Difference
	Mean [SD]	Mean [SD]	(Int.-Seg.) [t-stat]
Age	35.96	36.39	0.43
	[2.61]	[3.34]	[4.02]
% Graduated HS	73.42%	75.88%	-2.47%
	[10.73%]	[10.62%]	[-7.18]
% Graduated College	42.04%	36.38%	5.65%
	[13.85%]	[13.85%]	[3.78]
% Married	52.20%	59.53%	-7.32%
	[5.46%]	[6.79%]	[-3.39]
% working in FIN	6.91%	5.38%	1.54%
	[4.33%]	[3.46%]	[6.49]
Med HH Income	29,289	28,352	937.69
	[14986]	[12383]	[1.94]
Med Home Value	84,601	71,254	13,347
	[79,271]	[56,283]	[6.08]
Population Density	322.58	301.06	21.52
	[1933.60]	[2066.30]	[2.99]
Num. Counties	642	2,933	

Table 2: Summary Statistics: PSID Households

This table presents summary statistics for PSID households. Column (1) presents averages across households in the full PSID sample, column (2) presents averages for households that indicate residing in a county that is characterized as diverse as in eq. 2, and column (3) presents averages for households that indicate residing in a county that is characterized as racially segregated (i.e. not diverse). The sample consists of households for which there is available location data provided in the geocode file from the PSID. Variable definitions can be found in the appendix.

	Full Sample	Integrated	Segregated	Difference
	Mean [SD]	Mean [SD]	Mean [SD]	(Int.-Seg.) [t-stat]
% Black	14.37% [0.351]	16.10% [0.367]	13.61% [0.343]	2.49% [4.67]
% Minority	20.30% [0.402]	25.37% [0.435]	18.04% [0.385]	7.33% [5.78]
Age	50.2 [17.9]	49.6 [17.5]	50.5 [18.0]	-0.90 [3.25]
% Graduated HS	78.13% [0.41]	77.49% [0.42]	78.43% [0.41]	-0.94% [3.41]
% Graduated College	32.76% [0.469]	35.36% [0.478]	31.61% [0.465]	3.74% [2.54]
% Married	50.86% [0.500]	51.30% [0.500]	50.67% [0.500]	0.63% [5.97]
Income	33,959 [71,503]	37,870 [80,149]	32,223 [67,236]	5,646 [1.11]
# of Households	23,931	9,397	18,564	

Table 3: Summary Statistics for Brokerage Households

The sample is restricted to households with non-missing demographic information and to those households holding at least $1,000 in their portfolio at the end of 1991. Stock holdings are restricted to long positions in common stocks. Column one presents summary statistics for the full sample, and columns two and three break down these statistics based on whether the household resides in an integrated or segregated county as defined by equation 2.

	Full Sample	Integrated	Segregated	Difference
	Mean [SD]	Mean [SD]	Mean [SD]	(Int.-Seg.) [t-stat]
Portfolio Value	40,474 [145,714]	41,902 [159,458]	39,606 [136,671]	2,297 [9.86]
Num. Stocks	3.68 [4.46]	3.57 [4.30]	3.75 [4.55]	-0.185 [25.94]
Num. Trades	2.45 [3.52]	2.43 [3.11]	2.47 [3.75]	-0.040 [3.25]
% of HH with income<$50,000	35.9% [0.05]	32.6% [0.05]	38.0% [0.49]	-5.4% [-70.28]
% of HH with income between $50,000 and $100,000	40.1% [0.05]	42.1% [0.05]	38.9% [0.05]	-3.2% [40.47]
% of HH with income>$100,000	24.0% [0.04]	25.3% [0.04]	23.1% [0.04]	0.022 [32.45]
Num. Households	39,296	14,882	24,490	

Table 4: Baseline Effects of Integration on the Participation Decision

This table presents the results from linear probability models with standard errors clustered at the county level. The dependent variable is participation, which is an indicator variable taking the value of one if the household indicates holding publicly traded securities (excluding those held in retirement accounts) in a given survey year. The independent variable of interest, 'Integrated', is an indicator variable taking the value of one if the household resides in a diverse county as defined by equation 2. Residential location is time-varying and updated for each panel of the survey. Columns (2) and (3) add community ownership (the fraction of households in the county that participate excluding individual i) and connectedness (a scale of connectedness to help provided by the PSID) respectively. The other independent variables presented are individual level controls which include the following: age, square of age, interaction of gender and filing status, indicator for whether individual is Black, indicator for whether individual holds a college degree, income risk and the degree of risk aversion (a scale of risk avoidant behavior provided by the PSID). Untabulated individual controls include indicators for each percentile of the wealth distribution with the 50th percentile as the base category as well as indicators for every fifth percentile of the income distribution. All specifications also include a host of county-level controls and state-by-year fixed effects. See the appendix for a full list of controls used. ***, **, and * indicate significance at the 1, 5 and 10 percent level respectively

Dep Var: Participation	(1)	(2)	(3)
Integrated	0.010**	0.009**	0.012**
	[0.005]	[0.005]	[0.005]
Community Stock Ownership		0.041**	0.033**
		[0.016]	[0.016]
Individual Sociability			0.001
			[0.002]
Age	0.003***	0.003***	0.003***
	[0.000]	[0.000]	[0.000]
Age sq.	-0.005***	-0.002***	-0.002***
	[0.001]	[0.001]	[0.001]
Male x Married	0.063	0.066	0.072*
	[0.041]	[0.041]	[0.042]
Male	-0.003	-0.003	-0.004
	[0.005]	[0.005]	[0.005]
Married	-0.061	-0.062	-0.062
	[0.040]	[0.040]	[0.041]
Black	-0.068***	-0.067***	-0.070***
	[0.006]	[0.006]	[0.006]
College	0.109***	0.109***	0.109***
	[0.006]	[0.006]	[0.007]
Income Risk	-0.001	-0.001	-0.001
	[0.001]	[0.001]	[0.001]
Risk Aversion	-0.028***	-0.028***	-0.028***
	[0.005]	[0.005]	[0.005]
Income Indicators	Yes	Yes	Yes
Wealth Indicators	Yes	Yes	Yes
County Controls	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes
N	65,677	65,677	60,512
Adj. R-sq	0.320	0.320	0.323

Table 5: Household Relocation and the Effect of Integration

The dependent variable in all specifications below is participation which takes the value of one if the household indicates owning stock in a publicly traded column in year$t$. Column (1) evaluates the participation decision in the survey year that the household indicates moving from a racially segregated to a racially integrated county. Individual controls are from the year in which participation is measured. Income and income risk are first differences to control for any changes in income determining the participation decision at time t. County controls are at the destination level. Column (2) evaluates the participation decision in the survey year following the move to a diverse county. Since the PSID surveys are conducted every two years, this corresponds to two years after the move. Income controls are second differences to control for changes taking place between immediately prior to the move up to the period following the move. Column (3) evaluates the participation decision four years following the move to a diverse county and column (4) shows the effect of moving to diversity after six years. ***, **, and * indicate significance at the 1, 5 and 10 percent level respectively.

Dep Var: Participation	Year of Move	Two Years Post Move	Four Years Post Move	Six Years Post Move
	(1)	(2)	(3)	(4)
Moved to Integration$_t$	0.021***	0.023*	0.024*	0.023
	[0.011]	[0.011]	[0.012]	[0.014]
Community Stock Ownership	0.022	0.022	0.017	0.016
	[0.024]	[0.024]	[0.026]	[0.028]
Age	-0.000*	-0.000*	-0.000*	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Age sq.	0.002***	0.002***	0.003***	0.003***
	[0.001]	[0.001]	[0.001]	[0.001]
Male x Married	0.001	0.001	0.001	0.000
	[0.002]	[0.002]	[0.002]	[0.002]
Male	-0.011	-0.011	-0.016*	-0.025**
	[0.0080]	[0.008]	0.008]	[0.009]
Married	-0.052	-0.053	-0.055	-0.061
	[0.060]	[0.060]	[0.063]	[0.064]
Black	-0.063***	-0.063***	-0.063***	-0.059***
	[0.008]	[0.008]	[0.008]	[0.009]
College	0.093***	0.093***	0.090***	0.092***
	[0.008]	[0.008]	[0.009]	[0.010]
ΔIncome	0.001**	0.001**	0.001	0.001
	[0.000]	[0.000]	[0.000]	[0.001]
ΔIncome Risk	-0.001	-0.001	-0.001	-0.001
	[0.001]	[0.001]	[0.001]	[0.002]
Risk Aversion	-0.048***	-0.048***	-0.049***	-0.050***
	[0.008]	[0.008]	[0.008]	[0.009]
Income Indicators	Yes	Yes	Yes	Yes
Wealth Indicators	Yes	Yes	Yes	Yes
County Controls	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes
N	36,697	36,697	32,181	27,234
Adj. R-sq	0.346	0.346	0.354	0.362

Table 6: Instrumenting Integration

Panel A presents summary statistics for Great Migration Commuting zones broken down by integration classification. The statistics presented here are taken from the US decennial census and averaged over the PSID sample period (i.e. 1980-2010). A commuting zone is deemed to be integrated if the predicted Black population shock falls within the 40th and 60th percentile of the predicted shocks. CZs are classified as being segregated if it falls outside the 40th and 60th percentile of predicted shocks. Panel B presents the first and second stage results for 2SLS estimation where contemporaneous county integration is instrumented for using predicted Black population shocks in northern and western commuting zones. Data for the quantiles of predicted Black population shocks are taken from Derenoncourt (2019). Predicted shocks that fall within the 40th and 60th percentiles of the sample are defined as being diverse. All specifications include individual and county-level controls as in the baseline specification. ***, **, and * indicate significance at the 1, 5 and 10 percent level respectively.

Panel A: Commuting Zone Summary Statistics			
	Integrated	Segregated	Difference
	Mean [SD]	Mean [SD]	(Int.-Seg.) [t-stat]
Age	36.39	36.12	0.27
	[1.929]	[2.145]	[4.43]
% HS	77.43%	83.78%	6.35%
	[0.077])	[0.078]	[17.40]
% College	34.16%	29.15%	5.01%
	[0.004]	[0.002]	[12.02]
% in Finance	5.34%	5.25%	0.09%
	[0.023]	[0.020]	[34.87]
% Married	58.13%	58.60%	-0.47%
	[0.058]	[0.057]	[-16.02]
Med HH Income	32,475	34,011	-1,536
	[5,474]	[8,971]	[-4.14]
Med Home Value	84,229	95,468	-11,239
	[42,774]	[60,395]	[-27.82]
# of CZs	27	103	
Panel B: Two Stage Least Squares			
		First Stage	Second Stage
Pred. Population Shock		0.236***	
		[0.006]	
$\widehat{Integrated}$			0.056**
			[0.023]
Individual Controls		Yes	Yes
County Controls		Yes	Yes
State x Year FE		Yes	Yes
N		47,272	47,272
F-stat		169.5	–
Adj. R-sq		0.262	0.263

Table 7: Integration and Sociability

This table presents the results from several regression specifications where the integration indicator is interacted with several county-level characteristics. The dependent variable for all specifications is an indicator that takes the value of one if the individual reports owning equity. Column (1) interacts the integration indicator with an indicator for whether the county is in the top quartile of the population density distribution. Column (2) interacts the integration indicator with the number of civic organizations in a county. Column (3) interacts the integration indicator with the number of fitness centers/clubs in the county. All specifications include individual and county controls as well as state-by-year fixed effects. Standard errors are clustered at the county level. ***, **, and * indicate significance at the 1, 5 and 10 percent level respectively.

Dep. Var: Participation	(1)	(2)	(3)
Integrated × Pop.Density	0.040**		
	[0.019]		
Pop. Density	0.015		
	[0.010]		
Integrated × Civic Orgs		0.086*	
		[0.049]	
Civic Orgs		-0.039	
		[0.025]	
Integrated × Fitness Centers			0.204**
			[0.001]
Fitness Centers			-0.082
			[0.054]
Integrated	0.010*	0.005*	0.003
	[0.006]	[0.006]	[0.001]
Community Ownership	0.025	0.027	0.028
	[0.019]	[0.019]	[0.020]
Urban Ind.	0.017*	0.013	0.007
	[0.005]	[0.010]	[0.005]
Individual Controls	Yes	Yes	Yes
County Controls	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes
N	60,512	60,512	60,512
Adj. R-sq	0.332	0.331	0.332

Table 8: Alternative Definition of Integration: Spectral Segregation Index

This table presents the results from linear probability models that regress an indicator for participation on different variations of a spectral segregation index (SSI) developed by Echenique and Fryer Jr (2007). This index measures the degree of racial isolation of individuals across MSAs based on their friendship networks. The index can be calculated with respect to all races (i.e. the average isolation of White individuals, Black individuals, Hispanic individuals, etc.) as well as by individual race. Column (1) presents the results from the overall SSI, column (2) presents the SSI based on White individuals and column (3) presents the SSI based on Black individuals. All specifications include the same controls as Column (3) of the baseline specification. Standard errors are clustered at the MSA level. ***, **, and * indicate significance at the 1, 5 and 10 percent level respectively.

Dep. Var: Participation	Overall (1)	White (2)	Black (3)
Segregation	-0.029**	-0.018*	-0.012*
	[0.014]	[0.009]	[0.007]
Age	-0.001***	-0.001***	-0.001***
	[0.000]	[0.000]	[0.000]
Age sq.	0.003***	0.003***	0.003***
	[0.001]	[0.001]	[0.001]
Male x Married	0.003	0.002	0.002
	[0.002]	[0.002]	[0.002]
Male	-0.004	-0.004	-0.002
	[0.007]	[0.007]	[0.008]
Married	-0.081	-0.081	-0.092
	[0.070]	[0.070]	[0.072]
Black	-0.088***	-0.088***	-0.093***
	[0.008]	[0.008]	[0.009]
College	0.108***	0.108***	0.108***
	[0.009]	[0.009]	[0.010]
Income Risk	-0.003**	-0.003**	-0.002
	[0.001]	[0.001]	[0.001]
Risk Aversion	-0.042***	-0.042***	-0.036***
	[0.008]	[0.008]	[0.008]
County Controls	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes
N	24,823	24,823	22,250
adj. R-sq	0.338	0.338	0.350

Table 9: Portfolio Performance of Integrated and Segregated

This table presents summary statistics of household portfolio returns along with the results of regression specifications where the dependent variable is the excess value-weighted monthly (panel B) and the excess value-weighted monthly returns of a portfolio comprised only of local stocks (panel C). The local Sharpe ratio is computed by taking the excess local portfolio return and dividing by the standard deviation of the local portfolio. Holdings are considered to be local if the firm is headquartered within 250 miles of the household. Panel A also presents tests for differences in the mean performance measures between diverse and segregated households. All regressions in panels B and C include individual controls available in the brokerage data set as well as state-by-year fixed effects. These controls include age, gender, income category, marital status, children indicator, and home ownership indicator. Returns are winsorized at the 0.5% level to reduce the effects of small stocks. Standard errors are clustered by zip code.

Panel A: Statistics of Overall and Local Portfolio Performance

	$R_i - R_f$	$Sharpe$	$R_{local,i} - R_f$	$Sharpe_{local}$
	Mean (%) [SD]	Mean (%) [SD]	Mean (%) [SD]	Mean (%) [SD]
Integrated	1.383	0.308	1.471	0.158
	[0.012]	[0.052]	[0.021]	[0.002]
Segregated	1.353	0.286	1.319	0.151
	[0.009]	[0.060]	[0.017]	[0.001]
Difference ($Diverse-$ $Segregated$)	0.029**	-0.022	0.154***	0.006***
	[0.015]	[0.081]	[0.026]	[0.002]

Panel B: Regressions with $R_i - R_f$ as Dependent Variable

	(1)	(2)	(3)	(4)
Integrated	0.024	0.024	0.024	0.025
	[0.022]	[0.022]	[0.022]	[0.022]
$R_m - R_f$		1.069**	1.052**	1.090**
		[0.005]	[0.004]	[0.005]
SMB			0.470**	0.435**
			[0.007]	[0.007]
HML			0.086**	0.073**
			[0.005]	[0.005]
MOM				-0.226**
				[0.004]
Alpha	1.378**	0.270	0.106	0.281
	[0.065]	[0.065]	[0.065]	[0.065]
Individual Controls	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes
Number of Households	28,107	28,107	28,107	28,107
Number of Months	71	71	71	71
R-sq	0.014	0.125	0.139	0.143

Panel C: Regressions with $R_{local,i} - R_f$ as Dependent Variable

	(1)	(2)	(3)	(4)	(5)
Integrated	0.125**	0.123**	0.125**	0.125**	0.062*
	[0.040]	[0.039]	[0.039]	[0.039]	[0.037]
$R_m - R_f$		1.150**	1.121**	1.168**	0.205**
		[0.010]	[0.009]	[0.010]	[0.022]
SMB			0.612**	0.575**	0.502**
			[0.013]	[0.013]	[0.010]
HML			0.089**	0.078**	0.102**
			[0.011]	[0.011]	[0.010]
MOM				-0.267**	-0.257**
				[0.008]	[0.008]
$R_z - R_f$					0.915**
					[0.020]
Alpha	1.325**	0.163	-0.017	0.193	-0.675
	[0.120]	[0.119]	[0.119]	[0.119]	[0.118]
Individual Controls	Yes	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes	Yes
Number of Households	17,053	17,053	17,053	17,053	17,053
Number of Months	71	71	71	71	71
R-sq	0.013	0.095	0.111	0.115	0.136

Table 10: Trading Behavior Among Integrated and Segregated Investors

This table presents the results from regression specifications analysing the role of integration on various measures of trading behavior. All regressions are run at the household-month level and include state-by-year fixed effects with standard errors clustered at the zip code level. The dependent variable in column (1) is the number of stocks held in an investor's portfolio in a given month. The dependent variable in column (2) is the normalized portfolio variance, where normalized variance is computed as the standard deviation of the portfolio divided by the average standard deviation of stocks held in the portfolio. The covariance matrix is estimated using the past five years of monthly returns. The dependent variable in Column(3) average portfolio turnover, where average turnover is the average of purchase and sale turnover. The dependent variable in column (4) is the disposition effect, which is computed as the difference between the proportion of gains realized and the proportion of losses realized for each investor in a given month. The dependent variable in column (5) is the fraction of each household's portfolio that is local (firms headquartered within 250 miles of the household) adjusted for the fraction of the total market that is local.

	(1) # of Stocks Held	(2) Normalized Variance	(3) Avg. Turnover	(4) Disposition Effect	(5) % of Portfolio within 250m - % of Mkt within 250m
Integrated	-0.161**	0.014**	0.008*	0.001	0.044**
	[0.060]	[0.005]	[0.005]	[0.004]	[0.006]
Income	0.053**	-0.001	-0.003**	-0.003**	0.002*
	[0.015]	[0.001]	[0.001]	[0.001]	[0.001]
Age	0.026**	-0.001**	-0.002**	-0.000**	-0.001**
	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
Married	-0.175**	-0.004	0.008	0.003	0.014**
	[0.073]	[0.007]	[0.005]	[0.005]	[0.005]
Home Owner	0.166	-0.018	-0.013	0.008	0.021*
	[0.151]	[0.017]	[0.012]	[0.010]	[0.011]
Male	0.224**	-0.021**	-0.001	0.016**	-0.007
	[0.082]	[0.008]	[0.008]	[0.007]	[0.007]
Constant	1.925**	0.687**	0.390**	0.044**	0.128**
	[0.192]	[0.021]	[0.017]	[0.013]	[0.014]
State x Year FE	Yes	Yes	Yes	Yes	Yes
Number of Households	39,521	39,521	39,521	39,521	39,521
R-sq	0.022	0.005	0.044	0.009	0.048

Appendix

Table A1: Variable Definitions

Variable Definitions	
Geographic	
County Exposure	The probability that a Black individual will interact with a White individual in county j. Computed as in equation 1
Diverse	Takes the value of 1 for county j if county exposure is within 1/2 standard deviation of 50% exposure, 0 otherwise
Household level from PSID	
Participate	indicator taking the value of 1 if household i indicates owning shares of stock in publicly traded companies, mutual funds or investment trusts
Market Entry	indicator taking the value of 1 if the household enters the market having not participated in the prior survey
Minority	indicator for whether individual is a minority
Black	indicator for whether individual is Black
Age	age of individual
Age sq.	Square of age
Male x Married	interaction of gender and married
Income	Indicators for every fifth percentile of the income distribution.
Income Risk	standard deviation of the change in income over a five-year rolling basis.
Risk Aversion	indicator taking the value of one if individual scores higher than four points on the risk avoidance scale from the PSID
College	indicator for whether individual completed college
County level from census: Socioeconomic control variables	
ln(Median Home Value)	natural log of median home value
ln(Median Income)	percentile of median household income
Urban	urban/rural indicator
County Population Percentile	percentile of county population. Run specifications using an indicator for each percentile (0-99) with the 50th percentile as the omitted category.
% Married	proportion of married individuals
% HS	proportion of individuals that completed high school
% College	proportion of individuals that completed college
Avg. Age in County	average age in county j
County level from CRSP: Financial market control variables	
Public Firm Ind.	indicator taking the value of 1 if there exists a publicly traded firm headquartered in county j
% County Mkt Val	fraction of total US market value represented by local publicly traded firms
% Top Performers	fraction of local firms that were in the top 10% of nationwide performers that year
% Dividend Payers	fraction of local firms that pay dividends
Value-Weighted County Return	value weighted return of firms in county j
Top	indicator for whether the county is in the top decile of value weighted returns
Other	
sk	county level social capital index from the University of Pennsylvania
GM	percentile of predicted Black population shock in norther commuting zones from Derenoncourt (2019)
IV_dummy	takes the value of 1 if GM is within the 40th and 60th percentile

Table A2: Probit Model Specification

This table presents the results from probit model specifications that mirror the linear probability model specifications in tables 4 and 5.

Dep Var: Participation	Baseline (1)	Year of Move (2)	Two Years Post Move (3)	Four Years Post Move (4)	Six Years Post Move (5)
Integrated	0.045*				
	[0.026]				
Moved to Int.$_t$		0.129***	0.141***	0.100**	0.100
		[0.041]	[0.046]	[0.050]	[0.055]
Age	0.017***	0.017***	0.017***	-0.003**	-0.002
	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
Age sq.	-0.006**	-0.006	-0.002	0.013***	0.014***
	[0.003]	[0.003]	[0.004]	[0.005]	[0.005]
Male x Married	0.125***	0.117***	0.127***	0.010	-0.000
	[0.008]	[0.010]	[0.011]	[0.014]	[0.016]
Black	-0.831***	-0.836***	-0.851***	-0.427***	-0.420***
	[0.036]	[0.042]	[0.045]	[0.048]	[0.054]
College	0.744***	0.711***	0.702***	0.421***	0.418***
	[0.023]	[0.028]	[0.031]	[0.032]	[0.035]
Income Risk	0.016***	0.023***	0.022***	-0.014**	-0.015*
	[0.004]	[0.005]	[0.006]	[0.007]	[0.008]
Risk Aversion	-0.236***	-0.267***	-0.281***	-0.275***	-0.293***
	[0.025]	[0.032]	[0.036]	[0.039]	[0.044]
N	65,677	36,697	36,697	32,181	27,234
pseudo R-sq	0.221	0.216	0.221	0.367	0.366