Hispanic Culture, Portfolio Decisions, and Asset Prices*

Jawad M. Addoum, Carina Cuculiza, Alok Kumar, Stuart Webb

August 24, 2020

Abstract – We examine the impact of culture on portfolio decisions and asset returns, focusing on the large and growing Hispanic population in the United States. We find that both retail and institutional investors in high Hispanic neighborhoods overweight local, lottery-type, and high-momentum stocks and their trades are more strongly correlated. Systematic Hispanic trading generates excess return comovement among locally headquartered firms. We also find evidence of higher local stock returns, a stronger negative lottery-stock premium, and larger momentum returns in high-Hispanic areas. Collectively, these findings suggest that evolving ethnic demographics affect U.S. capital markets.

Keywords: Hispanic culture, demographic shifts, local bias, lottery stocks, return comovement, local returns, return momentum.

JEL classification: G11, G12.

*Please address all correspondence to Carina Cuculiza, Department of Finance, Miami Herbert Business School, 5250 University Drive, Coral Gables, FL 33124; *Email:* c.cuculiza@umiami.edu. Jawad M. Addoum is at S.C. Johnson College of Business, Charles H. Dyson School of Applied Economics and Management, Cornell University, 301B Warren Hall, Ithaca, NY 14853; *Email:* jaddoum@cornell.edu. Alok Kumar is at the Department of Finance, Miami Herbert Business School, 5250 University Drive, Coral Gables, FL 33124; *Email:* akumar@miami.edu. Stuart Webb is at the Carlson School of Management Finance Department, University of Minnesota, 321 19th Avenue South, Minneapolis, MN 55455; *Email:* swebb@umn.edu. We thank Wolfgang Ausseneg, Nick Barberis, William Bazley, Tim Burch, Vicki Bogan, Indraneel Chakraborty, Ngina Chiteji, Byoung-Hyoun Hwang, Narasimhan Jegadeesh, George Korniotis, Daniel Perez, Sheridan Titman, Andreanne Tremblay-Simard, Scott Yonker, seminar participants at the University of Miami, conference participants at the 2017 Boulder Summer Conference on Consumer Financial Decision Making, the 2017 Behavioural Finance Working Group Conference, the 2018 AFA Ph.D. Poster Session, the 2018 MFA Annual Meeting, the 2018 SWFA Annual Meeting, the 14th Annual Whitebox Advisors Graduate Student Conference at Yale University, and the American Economic Association (AEA) Committee on the Status of Minority Groups in the Economics Profession (CSMGEP) at the 2020 ASSA meetings for helpful comments and valuable suggestions. We are responsible for all remaining errors and omissions. An earlier version of the paper circulated with the title "Hispanic Culture, Stock Preferences, and Asset Prices."

Hispanic Culture, Portfolio Decisions, and Asset Prices

Abstract – We examine the impact of culture on portfolio decisions and asset returns, focusing on the large and growing Hispanic population in the United States. We find that both retail and institutional investors in high Hispanic neighborhoods overweight local, lottery-type, and high-momentum stocks and their trades are more strongly correlated. Systematic Hispanic trading generate excess return comovement among locally headquartered firms. We also find evidence of higher local stock returns, stronger negative lottery-stock premium, and larger momentum returns in high-Hispanic areas. Collectively, these findings suggest that evolving ethnic demographics affect U.S. capital markets.

Keywords: Hispanic culture, demographic shifts, local bias, lottery stocks, return comovement, local returns, return momentum.

JEL classification: G11, G12.

1 Introduction

Hispanics are the largest and fastest growing subculture within the United States. The 2015 Census Bureau estimates suggest that there are over 56.6 million Hispanic residents in the U.S. By 2060, the Hispanic population is projected to grow to 120 million individuals, constituting almost 30% of the U.S. population. At the same time, Hispanic households are growing wealthier. For example, between 2013 and 2016, the wealth gap between Hispanic and White households shrunk by almost 20%. Further, the Hispanic population is among the fastest growing in homeownership rates, entrepreneurship, income growth, and financial wealth. In particular, the percentage of Hispanic households earning income above $200,000 grew by over 80% between 2011 and 2017 and the number of Hispanic millionaires more than doubled between 2013 and 2016 (Hispanic Wealth Project, 2019).

Motivated by these broad demographic and cultural shifts in the U.S. population, in this paper, we examine whether the influx of Hispanic immigrants influence various financial market outcomes. In particular, we study the potential impact of Hispanic culture on retail and institutional investment decisions as well as asset prices. Our broad conjecture is that Hispanic investors' cultural background would affect their portfolio decisions. Given the size and rapid growth of the U.S. Hispanic population, their systematic portfolio decisions also have the potential to influence asset prices. These hypotheses build on two strands of literature.

First, our hypothesis that Hispanic culture may drive differences in portfolio decisions is motivated by the growing economics literature that examines how culture shapes economic decisions in general.[1] These papers suggest that various aspects of culture are important determinants of consumption, savings, and portfolio decisions.[2]

[1] Becker (1996) argues that "culture exercises a sizable influence over preferences and individual behavior" because it is "largely "given" to individuals throughout their lifetimes." In particular, he argues that "individuals have less control over their culture than over other social capital" since "they cannot alter their ethnicity, race, or family history, and only with difficulty can they change their country or religion." Further, Becker asserts that it is difficult to change culture and that culture exhibits a low depreciation rate, suggesting that the effects of culture may be long-lived and passed on through generations.

[2] Guiso, Sapienza, and Zingales (2006) provide a broad survey of the literature on culture and economics. For example, Guiso, Sapienza, and Zingales (2003, 2004a,b, 2006, 2008) demonstrate that nationality affects trust, an implicit requirement for economic exchange. Similarly, the nationality of American-born individuals' ancestors can affect women's labor force participation (Fernández, Fogli, and Olivetti, 2004) and fertility decisions (Fernandez and Fogli, 2009). Furthermore, Karolyi (2016) provides a more recent review of the literature on culture in finance. For instance, there is substantial evidence that both religion and language can affect investment style and preferences. Kumar, Page, and Spalt (2011) show that the predominant religious beliefs and gambling attitudes of a region can influence an institution's propensity to invest in certain types of stocks. Grinblatt and

Our second hypothesis, that Hispanics' distinctive stock preferences could affect asset prices and returns, is motivated by the finance literature that studies how demographic patterns affect aggregate stock returns.[3] Recent evidence in this literature suggests that differences in the underlying demographic structure and cohort sizes of the population can impact the demand for various goods, generating predictable patterns in financial asset returns. Despite providing evidence on the importance of generational cohorts, this literature has not yet examined how changes in the ethnic composition of the U.S. population might affect asset markets.

Understanding how specific ethnic and cultural groups' financial decisions impact asset prices is an interesting and important research question, but there are several empirical challenges. First, in order for researchers to identify distinct effects associated with an ethnic subgroup of interest, its members would not only have to exhibit decisions that differ from the local benchmark (e.g., White Americans), but each minority subgroup would also need to have preferences that sufficiently vary from one another. Second, uncovering asset pricing effects would necessitate that an ethnic subgroup has sufficient size and aggregate wealth such that its preferences could be plausibly reflected in prices.[4]

From this perspective, the Hispanic population is an ideal group of study, in that the Hispanic culture exhibits several unique features and consumption patterns (Korzenny and Korzenny, 2011; Cartagena, 2013).[5] In addition, Hispanics' prevalence at both the national and

---

Keloharju (2001) find that investors in Finland tend to overweight firms that disclose annual reports in their native language.

[3]Bakshi and Chen (1994) investigate the life-cycle investment hypothesis and the life-cycle risk aversion hypothesis and show that changes in the age structure of the population can affect the aggregate demand for financial investments and the equilibrium risk premiums. Poterba (2001), Abel (2003), and Ang and Maddaloni (2005) study the "baby boom" cohort and their effects on capital markets. Similarly, Goyal (2004) uses an overlapping generations framework to examine the relation between the population age structure, net outflows of the stock market, and stock market returns. He finds that market outflows are positively (negatively) correlated with the fraction of old (middle-aged) people. This evidence provides support for traditional life-cycle models. DellaVigna and Pollet (2007) show that cohort size fluctuations produce forecastable demand changes for certain sectors, which lead to predictable patterns in industry returns.

[4]Another prominent subculture within the U.S. is Chinese. However, due to data limitations, the sample is insufficient for our empirical tests. We collect data on Asian immigrants to the U.S. from the decennial census. The country of origin data sets for the Asian population are only available for the 2000 and 2010 decennial censuses. Since the time frame for our brokerage data is from 1991 to 1996, we are unable to test the stock preference conjectures on the Chinese subculture. Likewise, given the short availability of the Asian origin data sets, it is difficult to test the asset pricing effects. This issue is exacerbated by the fact that even if the Chinese population in the U.S. is growing, it is still relatively small compared to other subgroups. The 2010 statistics show that the Hispanic population in the U.S. constitutes 16.3% of the total population, while the Chinese population only constitutes 1.23%. *Sources:* https://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf; https://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_SF1_PCT7&prodType=table

[5]Cartagena (2013) studies the unique characteristics of Hispanic consumers to learn how U.S. businesses can

especially regional levels is likely to provide the impact necessary to uncover asset pricing effects associated with their decisions.[6]

We begin our empirical analysis by examining the revealed stock preferences of Hispanic investors and document several distinctive features of their portfolios. To guide our approach, we make specific conjectures on how Hispanic investors' investment preferences are likely to differ from those of the typical U.S. investor. These conjectures are grounded in salient dimensions of the Hispanic culture.

For instance, Hispanic individuals generally emigrate from countries with institutions that are relatively corrupt in comparison to those in the U.S. (Transparency International, 2007). This exposure to corruption is likely to generate distrust toward financial institutions that may persist even after moving to the U.S. (Guiso, Sapienza, and Zingales, 2004b, 2006; Osili and Paulson, 2008).[7] In turn, Hispanics' lower trust levels have the potential to affect their portfolio decisions, translating to a lower propensity to invest in the stock market (Osili and Paulson, 2008). However, Guiso, Sapienza, and Zingales (2008) show that knowledge about local institutions can counteract mistrust, suggesting that Hispanic investors would exhibit a preference for geographically proximate investments that are more easily monitored. These insights motivate our first conjecture: Hispanic investors are likely to exhibit a stronger local bias than the average U.S. investor.

Hispanic individuals are also likely to exhibit stronger "Keeping-Up-with-the-Joneses" preferences (Gali, 1994; Soriano, 1995; Demarzo, Kaniel, and Kremer, 2004; Hong, Jiang, Wang, and Zhao, 2014). In particular, Korzenny and Korzenny (2011) argue that in Hispanic communities, being less wealthy than one's neighbors can generate significant disutility. As a result, status-seeking Hispanics may exhibit a stronger propensity to gamble to increase their personal wealth and move up the social ladder (Friedman and Savage, 1948; Brunk, 1981; Brenner, 1983; Becker, Murphy, Werning, et al., 2000). Consistent with this behavior, the Hispanic population is one of the largest buyers of lottery tickets (Clotfelter and Cook, 1990; Stranahan and Borg,

capture a larger portion of the Hispanic market. They tend to focus on salient features of the Hispanic culture, such as community leadership, tight social networks, and collectivism.

[6]Another advantage of focusing on the Hispanic population is that areas that are predominantly Hispanic tend to have a low concentration of other ethnicities. For instance, the correlation between high-Hispanic and high-Black counties is -0.008. This allows us to disentangle whether our effects are specific to the Hispanic culture versus being associated with residents of a particular area.

[7]For instance, Weaver (2003) shows that in comparison to non-Hispanic Americans, Mexican Americans have less trust in people running banks, financial institutions, and major businesses.

1998). From a portfolio choice perspective, these insights suggest that Hispanics who do not want to lag behind their peers would invest relatively heavily in lottery stocks (Kumar, 2009), leading to our second conjecture: Hispanics are likely to overweight lottery stocks relative to the average U.S. investor.

Another distinctive characteristic of Hispanic consumers is that they find commonly-used and branded products to be more appealing than one used by only a few (Korzenny and Korzenny, 2011). This aspect of culture also has potential implications for portfolio choices of Hispanic investors and asset prices in predominantly Hispanic areas. First, as others buy a stock, it becomes more and more desirable, which would generate herding among Hispanic investors (Hong, Kubik, and Stein, 2004). In turn, this coordinated buying may drive up prices, leading to higher future short-term returns and potential for subsequent return-chasing by Hispanic investors.[8] Together, Hispanic preferences for socially "certified" products and strong "Keeping-up-with-the-Joneses" preferences motivate our third conjecture: Hispanic investors are likely to overweight high-momentum stocks relative to the average U.S. investor.

We test these conjectures using institutional holdings data from the 13(f) Thomson Reuters database and retail portfolio holdings at a large U.S. discount brokerage house. Since we cannot directly observe each investor's ethnicity, we use the concentration of Hispanic residents in an investor's local area (i.e., county or ZIP code) as a proxy. While we acknowledge that this is an indirect approach to testing whether the Hispanic culture affects investor preferences, the extant literature suggests that local culture can influence the behavior of all market participants, including institutional and retail investors (Kumar, Page, and Spalt, 2011). We use the institutional and retail investor data sets to provide micro-level evidence that investors living in high-Hispanic areas are influenced by the Hispanic culture. Further, the systematic patterns we document motivate our later asset pricing tests.

Our results indicate that even after controlling for a large set of known determinants of stock selection, institutional and retail investors residing in predominantly Hispanic areas are significantly more likely to invest in local, lottery, and high-momentum firms. To establish a tighter relation between the Hispanic culture and investor preferences, we address potential

[8] Hong, Jiang, Wang, and Zhao (2014) show that "Keeping-Up-with-the-Joneses" preferences generate return chasing behavior. Specifically, when the market values of local stocks are high (low), the wealth of the peer group is high (low), and hence, there is a greater (lower) demand for local risky assets.

endogeneity concerns. Specifically, unobservable geographic heterogeneity may be correlated with both the prevalence of Hispanics in a given area and local residents' portfolio decisions.[9]

We address endogeneity concerns in two ways. First, we use the minimum distance from the Canadian border to each investor's location as an instrumental variable (IV) and find effects similar to those in our baseline stock preference regressions. Importantly, since the U.S. Hispanic population is concentrated closer to the southern border (see Figure 1), we find a strong first-stage relationship between our distance IV and the concentration of Hispanic residents in a local area.[10] Further, other than through preferences rooted in Hispanic culture, there are no clear channels through which investors who live farther from the Canadian border should prefer local, lottery, or high-momentum stocks.

In addition to this IV approach, we further control for sources of unobserved heterogeneity, focusing on financial sophistication, using the retail investor data set. In particular, we add MSA-level fixed effects and interact our Hispanic concentration variable with two measures of portfolio concentration (Rajan and Zingales, 1998).[11] We find that holding cross-sectional differences in financial sophistication constant, retail investors living in high-Hispanic localities continue to exhibit stronger local bias, overweight lottery stocks, and chase returns. The consistency in our estimates across the ordinary least squares (OLS), IV, and Rajan and Zingales (1998) approaches suggests that the effects of the Hispanic culture on portfolio decisions are unlikely to be driven by correlated unobservables.

Building on these results, we turn to the question of whether investors in high-Hispanic areas affect the stock prices and returns of local companies. Relative to the independent baseline culture in the U.S., the Hispanic population exhibits more "group-oriented" behavior, where people tend to conform with widely accepted social norms and avoid deviating from the consensus (Markus and Kitayama, 1991). As a result, Hispanic investors are likely to affect asset prices and returns by systematically trading (i.e., herding) in the same subset of stocks (Barber, Odean, and Zhu, 2009b).

We extend the results of our Hispanic stock preference tests and examine whether there is

[9]One particular concern is that investors who exhibit stronger return chasing and local bias tendencies may differ on the dimension of financial sophistication.

[10]Figure A1 plots the Hispanic concentration per county for the entire U.S., not only the counties where companies are headquartered.

[11]We use portfolio concentration as a measure of financial sophistication (Grinblatt and Keloharju, 2001).

stronger herding in local, lottery, and high-momentum stocks in high-Hispanic areas. Following Lakonishok, Shleifer, and Vishny (1992) and Barber, Odean, and Zhu (2009b), we create a herding measure that determines whether there is systematic buying pressure in a specific set of stocks. Consistent with our conjecture, we find evidence that institutional and retail investors residing in high-Hispanic areas have a stronger propensity to herd in local, lottery, and high-momentum firms.

In light of both our overweighting and herding results, we examine the asset pricing effects of these culture-driven preferences and systematic trading. First, we test whether the correlated trades of investors induce a common factor in the returns of local, lottery, and high-momentum stocks, leading to excess comovement that cannot be explained by firm fundamentals (Barberis, Shleifer, and Wurgler, 2005). Since Hispanic investors are more likely to herd in these stocks, we expect the overall comovement of returns to be higher in high-Hispanic areas. This hypothesis is also in line with Eun, Wang, and Xiao (2015), who suggest that in countries with culturally tight and "group-oriented" traits, stock prices experience higher comovement. We find that, on average, local, lottery, and high-momentum stocks of firms headquartered in high-Hispanic areas comove significantly more than those located in low-Hispanic areas.

Next, we perform three sets of cross-sectional sorts aimed at understanding whether Hispanic investors' tendency to overweight local, lottery, and high-momentum stocks generate predictable pricing and return patterns. Our first set of tests is based on the idea that in locales where Hispanics are especially prevalent, investors who overweight local stocks may potentially drive prices and expected returns (Hong, Kubik, and Stein, 2008). In particular, the greater demand for local assets can lead to higher average short-term returns.

To test this conjecture, we construct a trading strategy using a bivariate dependent sort. First, we sort stocks into quintiles based on the concentration of Hispanics in the Metropolitan Statistical Area (MSA) where each company is headquartered. Second, we sort stocks into quintiles based on their previous quarter change in local institutional ownership (e.g., institutional investors who are within a 60 mile radius from each company's headquarter). For each Hispanic quintile, we construct a zero-cost portfolio by going long in the top quintile and short in the bottom quintile based on previous quarter changes in local institutional ownership.

We find a positive and statistically significant Long-Short alpha in the high-Hispanic quintile, indicating that local investors' purchases (sales) generate higher (lower) abnormal short-term returns. In contrast, the zero-cost trading strategy does not generate an abnormal return in the low-Hispanic quintile. Importantly, the Long-Short return among stocks in the high-Hispanic quintile is significantly larger than that among stocks in the low-Hispanic quintile, reflecting Hispanic investors' overweighting of local stocks and consequent importance in their pricing. In further analyses we find that the difference in abnormal returns disappears in about eight months.

In our next cross-sectional tests, we examine the effect of Hispanic investors' preferences on the returns of lottery-type stocks. Theory and empirical evidence suggest that stocks with high idiosyncratic volatility, high idiosyncratic skewness, and low prices typically earn negative returns (Barberis and Huang, 2008; Kumar, Page, and Spalt, 2011). We investigate whether the magnitude of the negative lottery-stock premium is larger for the subset of companies headquartered in high-Hispanic areas.

As before, we test this conjecture using a bivariate dependent sort, where we first sort stocks into lottery and non-lottery classifications. Then, within each group, we sort firms into quintiles based on the local Hispanic concentration measure. Consistent with our conjecture, we find that the lottery-stock premium has larger magnitude (i.e., is more negative) among stocks with a higher local concentration of Hispanics. This trend cannot be explained by well-known risk factors. A zero-cost strategy that goes long lottery stocks in high-Hispanic areas and short lottery stocks in low Hispanic areas generates a risk-adjusted monthly return of -0.53% (t-statistic $= -2.17$). Conversely, there is a positive association between the returns of non-lottery stocks and the local concentration of Hispanic residents in an area.

Finally, we test whether momentum returns are stronger among firms headquartered in areas with a high concentration of Hispanics. Since investors in high-Hispanic areas overweight both high-momentum and local stocks, they are likely to chase returns among local stocks, potentially amplifying the momentum effect.[12] We test this conjecture using a double-sort

[12] Jegadeesh and Titman (1993) and De Long, Shleifer, Summers, and Waldmann (1990) suggest that the momentum phenomenon is consistent with trend chasing behavior. Jegadeesh and Titman (1993) argue that "positive feedback traders", e.g. those exhibiting trend chasing behavior, can cause momentum patterns in asset returns. Further, De Long, Shleifer, Summers, and Waldmann (1990) posit that by purchasing stocks with recent price increases, uninformed investors with extrapolative expectations can induce rational investors, hoping to profit from uninformed investors' continued purchase of increasing assets, to chase returns as well. The end result of this feedback is the well-documented momentum anomaly.

portfolio approach. First, we sort all companies into quintiles based on the local, MSA-level Hispanic concentration measure. We then sort all stocks within each quintile into winners and losers based on lagged 6-month formation period returns.

We find that the momentum portfolio of companies headquartered in high-Hispanic areas generates an average monthly return of 0.79% (t-statistic = 3.27). In contrast, the momentum portfolio formed using firms headquartered in low-Hispanic areas generates a statistically insignificant average monthly return. These momentum profits cannot be explained by known risk factors. Across various linear factor models, the alpha of the high-Hispanic momentum portfolio exceeds that of the low-Hispanic momentum portfolio by between 0.60% and 0.70% per month. This risk-adjusted performance differential is consistently significant at the 1%-level. Overall, our collective evidence suggests that known risk factors do not account for the culture-based geographic segmentation of the momentum effect.[13]

These findings contribute to several strands of the finance literature. First, we add to the growing literature that examines how culture affects financial decisions. Recent evidence shows that cross-country differences in the degree of collectivism versus individualism affect aggregate market outcomes. In particular, Chui, Titman, and Wei (2010) analyze 55 countries and find that momentum profits can be attributed to differences in individualism. However, Chui, Titman, and Wei (2010) also acknowledge that the momentum effect could be stronger in areas where individuals exhibit more "group-oriented" behavior. Specifically, they state that "investors in less individualistic cultures [e.g. "group-based" cultures] place too much credence on consensus opinions, and may thus exhibit herd-like overreaction to the conventional wisdom" (page 389). Consistent with this observation, we show that retail and institutional investors in high-Hispanic areas exhibit herding and that momentum profits in the U.S. are concentrated among firms located in high-Hispanic areas.[14]

[13]We perform several robustness tests to ensure that the stock preference results and asset pricing effects we document are driven by the preferences of investors living in high-Hispanic areas. First, we look at a second asset market, the U.S. residential real estate market, because it provides a tighter link between the buyers of an asset and the asset itself. The return chasing and preference for commonly used products in Hispanic communities implies that as more and more families buy real estate in areas with a high concentration of Hispanic residents, more people will find owning a house in this area desirable, potentially leading to price run-ups. Consistent with this intuition, we find that house prices in areas with a large Hispanic population experience higher realized volatility due to more pronounced price run-ups and subsequent downturns. We also re-perform our portfolio choice analyses and show that they are robust to: (i) using the minimum distance from the Mexican border as an alternative instrumental variable, (ii) controlling for investor-level financial knowledge and investment experience, and (iii) using the concentration of Hispanics relative to the total population of an area as an alternative measure.

[14]Similarly, Cheon and Lee (2017) study 42 countries and contend that investors in more individualistic coun-

By extension, our paper contributes to the literature examining the momentum return anomaly (Jegadeesh and Titman, 1993). Existing studies suggest that some determinants of momentum profits include lagged market returns (Cooper, Gutierrez, and Hameed, 2004; Stambaugh, Yu, and Yuan, 2012), investor sentiment (Antoniou, Doukas, and Subrahmanyam, 2010), market illiquidity (Avramov, Cheng, and Hameed, 2015), and macroeconomic factors (Liu and Zhang, 2008), among others. Again, to our knowledge, we are the first to demonstrate that "group-based," rather than only individualistic, cultural traits are also associated with pronounced momentum profits.

Our results also contribute to the literature on excess return comovement (Barberis, Shleifer, and Wurgler, 2005). Related to our work, Eun, Wang, and Xiao (2015) perform a cross-country analysis and find that in culturally tight and "group-oriented" countries, stock prices experience higher comovement. Our results are in line with theirs, but add to the collective understanding of comovement in several important ways. First, we show that there is significant geographic variation in comovement even within the U.S. Second, we distinctively focus on the connection between this cross-sectional variation in return comovement and the degree of local Hispanic population concentration, showing that stocks in high-Hispanic areas experience stronger comovement that cannot be explained by fundamentals. Third, our asset pricing tests are supported not only by conjectures about local investors' preferences, but by portfolio-level evidence of return chasing and herding behavior by both institutional and retail investors in high-Hispanic areas.

Finally, we advance the literature that investigates how changes in demographic patterns affect asset markets. Motivated by the aging baby boom generation, existing studies focus on how evolution of age structure in the population can generate systematic patterns in consumer demand and, in turn, affect asset returns (e.g. Bakshi and Chen, 1994; Goyal, 2004; DellaVigna and Pollet, 2007). An important contribution of our study is to shift the focus toward culture as another important source of variation in the population structure in the United States. Our focus on the large and growing Hispanic population demonstrates that changes in the ethnic composition of the U.S. population can also have significant effects on asset markets.

tries tend to be more overconfident and have a preference for high MAX stocks.[15] We complement their findings by showing that investors with stronger "Keeping-Up-with-the-Joneses" preferences tend to overweight lottery stocks.

2 Data and Summary Statistics

In this section, we briefly describe the data used in our empirical analysis. We use several data sources, including the 13(f) Thomson Reuters database, brokerage data from a large discount brokerage house, decennial census data from the United States Census Bureau, and data from the Center for Research in Security Prices (CRSP) and COMPUSTAT.

2.1 Institutional Holdings and Location Data

We collect quarterly common stock holdings of 13(f) institutions from the Thomson Reuters database. The sample period is from 1980 to 2018. We also obtain the location (i.e., ZIP code) for each institutional investors using the Nelson's Directory of Investment Managers, investigating the Securities and Exchange Commission (SEC) documents, and by searching the Web sites of institutional managers. Our main institutional investor sample excludes banks and insurance companies, as these are known to hold portfolios that are relatively more conservative (Kumar, Page, and Spalt, 2011).[16] Consequently, they could be less likely to exhibit preferences that are consistent with the Hispanic culture. Since the 13(f) Thomson Reuters database does not include institutional investors' trades, we complement this data with trading data from ANcerno Ltd., which is available from January 1, 1997 to December 31, 2010.

2.2 Household Brokerage and Demographic Data

To examine retail investor preferences, we use monthly retail investment account data from a large discount brokerage house. The main advantage of using this data set is that it includes household portfolio holdings and transactions at the ZIP code-level for 51,957 households from January 1991 to November 1996.[17] Panel A of Table 1 presents summary statistics for these brokerage data. Of the nearly 52,000 households, almost 74% are married and 97% own their own home. Consistent with the empirical fact that men are more likely to participate in the stock market, 88% of the heads of household in our sample are men. In addition, the median income of the households is $62,500, making them wealthier than the average household in the United States. In their portfolios, they tend to hold almost three securities with an average

[16]The results are qualitatively similar if we include all type of institutions.

[17]See Barber and Odean (2000) for further details on the brokerage data.

dollar value in each security of a little over \$9,000. Furthermore, the average Sharpe Ratio of these portfolios is 12.4%.

2.3 United States Decennial Census

We combine the institutional and brokerage data with data from the decennial censuses of the United States Census Bureau. For the institutional investor preferences tests, we use the decennial censuses from 1980 to 2010. To analyze retail investor preferences, we utilize the 1990 census. For the asset pricing tests, we use the decennial censuses from 1970 through 2010. Panel B of Table 1 presents summary statistics for all 29,305 ZIP codes in the 1990 decennial census of the U.S. Panel C restricts the sample to the 10,484 ZIP codes in which at least one household from our brokerage sample resides. Panel D of Table 1 reports summary statistics for all U.S. counties from 1980 to 2010. Panel E restricts the sample to the almost 300 counties in which at least one institutional investor is located.

The summary statistics show that brokerage customers and institutional investors tend to live in ZIP codes or counties that are more urban and populous. More specifically, they have an average population of 17,399 and 488,477, respectively, which are more than twice the average population of all U.S. ZIP codes and counties. Moreover, the ZIP codes in which our brokerage households reside are wealthier than the average ZIP code in the U.S., with a median income of \$62,000, closely matching our retail investor data. Finally, a higher percentage of residents in brokerage ZIP codes and institutional counties are Hispanic (6.1% in brokerage ZIP codes versus 4.4% in an average U.S. ZIP code and 7.8% in institutional counties versus 5.7% in an average U.S. county). Importantly, we find a similar Hispanic concentration in both brokerage ZIP codes and institutional counties, about 0.1.

2.4 Stock-Level Data

We also use the standard data sets when analyzing common stocks, CRSP and COMPUSTAT. From CRSP, we utilize monthly stock prices, returns, and shares outstanding from January 1970 through December 2018. We restrict our sample to include only common shares, using observations with share codes of 10 or 11. We obtain the location of each company's headquarters from the annual COMPUSTAT data files. In addition, we use the monthly Fama-French factors from Kenneth French's data library and the liquidity factor from Lubos Pastor's website.[18]

For our asset pricing tests, we use a longer time period than in the investor preference analyses. We use the institutional and retail investor data to motivate our asset pricing tests and to provide supporting micro-level evidence, which strengthens our overall conclusions.

2.5 Instrumental Variable Distance Measure

From the U.S. Gazetteer Files of the United States Census Bureau, we collect the latitude and longitude coordinates of the centroid of each ZIP code and county in the U.S. We then calculate the minimum distance from the Canadian border to each ZIP code and county using the Haversine formula (Vincenty, 1975; Coval and Moskowitz, 1999, 2001).

3 Hispanic Culture, Stock Preferences, and Comovements

In this section, we present our first set of main results. In particular, we analyze whether Hispanic investors overweight local, lottery, and high-momentum stocks. We implement an instrumental variable approach to mitigate the potential for unobservable variables biasing our ordinary least squares (OLS) results. We also investigate whether Hispanic investors' trades are correlated, consequently generating excess return comovement that cannot be explained by fundamentals.

3.1 Identifying Cultural Style Preferences

We examine whether Hispanics tend to overweight local, lottery, and high-momentum stocks. To identify the Hispanic population in the U.S., we use various decennial census data sets. Even though the group of people who identify themselves as Hispanics is diverse, the majority (79% according to the 1990 decennial census) of Hispanic individuals have origins in one of three different countries: Mexico, Puerto Rico, and Cuba.[19] Of the other countries typically mentioned in the decennial census, the majority of individuals come from Central America or

[18] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html; http://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2016.txt

[19] Approximately 62% of Hispanic Americans identify themselves as Mexicans in the 1990 decennial census. In addition, 12% claimed their country of origin to be Puerto Rico and almost 5% named Cuba.

the Dominican Republic. Nevertheless, every Hispanic country is represented.[20] We abstract from differentiating individuals based on their country of origin and instead treat the Hispanic population as a single subculture within the U.S. This approach is motivated by work on cultural clustering (e.g., Ronen and Shenkar (2013)) which shows that Spanish-speaking Latin American countries exhibit a high degree of cultural similarity.

We test our conjectures by regressing the excess weight of investor i's portfolio on the set of stocks, s, at time t normalized by the market weight of set s, that is $EW_{i,s,t} = (w_{i,s,t} - w_{m,s,t})/w_{m,s,t}$, on a vector of local area (i.e., county or ZIP code) and investor level variables. They analysis is at the county-level for institutional investors and at the ZIP code-level for retail investors. For the institutional-level analysis, we follow Kumar, Page, and Spalt (2011) and perform the analysis at the county-level, since some of the controls for the regressions are not available at the ZIP code level for the 1980 decennial census.

The set of stocks varies depending on the conjecture we are testing. For local stocks, s is the set of stocks headquartered within 60 miles of investor i's location. For lottery stocks, we follow Kumar (2009) and use three stock characteristics to identify stocks that might be perceived as lotteries: (i) stock specific or idiosyncratic volatility, (ii) idiosyncratic skewness, and (iii) stock price. Hence, s is the set of stocks that meets all of these criteria: being in the lowest 50$^{\text{th}}$ stock price percentile, the highest 50$^{\text{th}}$ idiosyncratic volatility percentile, and the highest 50$^{\text{th}}$ skewness percentile of the CRSP/Compustat sample. All three sorts are carried out independently.[21] However, as Kumar, Page, and Spalt (2011) suggest, some institutional investors might have constraints that prevent them from investing in very low priced stocks and; therefore, we do not use stock price as one of the lottery-stock attributes for the institutional-level preferences analyses.[22] For high-momentum stocks, s is the set of stocks in the top decile of returns over the most recent 12 and 6 month periods of the CRSP/Compustat sample.

Our main independent variable is H/W, which is defined as the census count of Hispanic individuals divided by the census count of White individuals in investor i's area (i.e., county or ZIP code). We scale the variable by the population of Whites because we want to compare how

[20] Following the U.S. Census, we define Hispanic countries as those that share a Spanish culture or origin. These include Cuba, Mexico, Puerto Rico, and countries in Central and South America. This classification excludes non-Spanish speaking countries in Latin America, such as Brazil.

[21] On average, 20% of the CRSP/Compustat sample are classified to be lottery stocks.

[22] In untabulated results, we find that the institutional-level results are broadly consistent with the main findings if we use all three conditions to define lottery stocks.

Hispanics invest relative to their White American peers. H/W measures the concentration of Hispanics in an area; thus, in localities where H/W is high, investors should overweight local, lottery, and high-momentum firms.[23]

3.1.1 Institutional Investor-Level Preferences

To analyze institutional-level investor preferences, we include several county-level demographic variables that are consistent with the literature (e.g., Kumar, Page, and Spalt (2011) and Hilary and Hui (2009)). Persons is the total population of the county. Urban is the proportion of the county that lives in an urban area. Male-female Ratio is the percentage of male to female residents in the county. Median Age is the median age of county residents. Married is the proportion of county households that are married. Education is the proportion of county-level population over the age of 25 with at least a bachelor's degree. We also include portfolio-level controls. In particular, Portfolio Value is the total market value of the investor's portfolio and HERF is the Herfindahl Index of the institution's portfolio. All non-indicator variables are standardized, and county-time clustered t-statistics are presented in parentheses below point estimates.

We augment the baseline regression specification with time (year-quarter) fixed effects to remove any time trends. We also include state-level fixed effects to remove any geographic, time invariant, unobservable variables that could be driving our results. By including the state-level fixed effects, we are comparing institutional investors within the same state but with different exposures to the Hispanic culture, depending on the county in which they are located.

We present estimates from our primary tests examining whether the Hispanic culture impacts institutional investors' portfolio decisions in Table 2, Panel A.[24] Column (1) shows that investors in high-Hispanic areas tend to have a stronger local bias. The coefficient on H/W is positive and statistically significant, 27.62 (t-statistic = 4.76). This effect is also economically significant, indicating that a one standard deviation increase in H/W is associated with a portfolio weight in local stocks that is almost 28% higher than the market weight.

Consistent with our conjecture, the evidence in column (2) suggests that the local Hispanic

[23]In the Internet Appendix, we show that the results of this section are robust to scaling the Hispanic population by the total population of an area.

[24]The univariate regressions can be found in Table A3. The results show that the H/W coefficient is positive and statistically significant, consistent with the specifications that include county and institutional level controls.

culture affects institutional investors' preferences, as the H/W coefficient is positive and statistically significant when the dependent variable is the excess weight on lottery stocks. The coefficient is also significant in economic terms. Specifically, a one standard deviation increase in H/W corresponds to a 0.03% increase in the weight assigned to lottery-type stocks.

The last two columns of Panel A show the results of regressions, where the dependent variables are the excess weight on stocks that have performed well the in the past 12 and 6 months, respectively. As hypothesized, we find that institutional investors who are located in high-Hispanic areas overweight high-momentum stocks. In particular, a one standard deviation increase in H/W translates to 0.10% (t-statistic = 6.60) higher weight in stocks that have experienced high returns in the past 12 months. Similarly, a one standard deviation increase in H/W is associated with a 0.07% increase in the weight assigned to stocks with high returns during the past 6 months.

3.1.2 Household-Level Preferences

For the retail-level tests, we include several ZIP code-level variables as controls. Persons is the census count of individuals in household i's ZIP code. B/W measures the concentration of Blacks relative to the White population, and is calculated as the census count of Blacks divided by the census count of White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. $\mathrm{I}_{\mathrm{Urban}}$ is an indicator variable taking the value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density and is calculated as the total population of the ZIP code divided by its land area. We include several household level controls, including indicators if the head of the household is male ($\mathrm{I}_{\mathrm{Male}}$), married ($\mathrm{I}_{\mathrm{Married}}$), and owns his home ($\mathrm{I}_{\mathrm{Own\ Home}}$). We also control for household i's income category, the age of the household, and the level of education. In addition, we include portfolio level controls, such as the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). All non-indicator variables are standardized, and ZIP code-year clustered t-statistics are presented in parentheses below point estimates.

We also augment our baseline household investment preference regressions with MSA-level fixed effects.[25] In this way, we are able to control for unobserved geographic heterogeneity at the MSA-level that could be correlated with the Hispanic concentration. By including MSA-level fixed effects, we are able to compare investors, who reside in the same MSA, but have different exposures to the Hispanic culture at the ZIP code-level, and thus, may have differing preferences regarding local, lottery, and high-momentum stocks.

We present the estimates for the retail preference tests in Table 2, Panel B.[26] Columns (1) and (2) suggest that culture may influence preferences for local stocks. Column (1) includes ZIP code and household level demographic controls, while column (2) also accounts for geographic heterogeneity at the MSA-level with MSA fixed effects. In column (1), the coefficient of H/W is large and statistically significant, showing that investors in high H/W ZIP codes invest relatively more in local companies compared to the market weight. The H/W coefficient is 253.81 (t-statistic = 8.21), indicating that a one standard deviation increase in H/W is associated with a portfolio weight in local stocks more than three times the market weight. In column (2), the H/W coefficient is statistically insignificant. However, this could be potentially driven by endogeneity, which we will address in the following section.

In columns (3) and (4), coefficients are presented where the dependent variable is the excess weight on lottery stocks. Column (4) includes MSA fixed effects for the MSA in which household i resides, while column (3) does not. In each regression, the estimate for H/W is statistically significant. More specifically, investors in high-Hispanic areas tend to overweight lottery stocks. For instance, the coefficients of 15.29 (t-statistic = 3.86) in column (3) and of 9.11 (t-statistic = 2.09) in column (4) suggest that a one standard deviation increase in H/W for household i's ZIP code translates to a 10% higher weight on lottery stocks. From the last three columns of Table 2 Panel B, we see that people who reside in high H/W ZIP codes tend to overweight stocks that have recently had high returns. The coefficients on H/W are statistically significant for the most recent 12 and 6 month returns with coefficients of 2.61 (t-statistic = 2.65) and 2.41 (t-statistic = 2.26), respectively.[27]

[25]For the retail investor data set, we do not augment our OLS and IV regressions with ZIP code-level fixed effects because H/W and MinDist are ZIP code-level variables that do not vary over time. Therefore, H/W and MinDist would be absorbed by these fixed effects.

[26]The univariate regressions can be found in Table A3. The results show that the H/W coefficient is positive and statistically significant, consistent with the specifications that include ZIP code and household level controls.

[27]Magnitudes are much smaller relative to the results for local bias due to the relatively small market weight of "local" stocks for most investors.

3.2 Addressing Endogeneity Concerns

Building on our baseline results in Table 2, we turn to addressing potential endogeneity concerns associated with omitted variables. [28] In particular, unobservable variables may bias our OLS regression coefficients because of unobserved heterogeneity across counties or ZIP codes that is correlated with an area's Hispanic concentration. The sign of the bias cannot be determined because there are likely to be competing forces regarding the nature of the endogeneity. An example of an omitted variable that may understate our OLS results is financial sophistication. Since high-Hispanic areas tend to be wealthier and more urban, investors living in these areas could be more financially sophisticated, and thus, less likely to exhibit a local bias, invest in lottery stocks, or chase returns. Similar to financial sophistication, there could be other unobservable variables that may over- or under-state our OLS regression estimates.

3.2.1 Instrumental Variable Approach

To address the unobservable variables concern, we use the minimum distance from the Canadian border (MinDist) to investor i's locality (i.e., county or ZIP code) as an instrumental variable.[29] We create the MinDist variable by using the latitudes and longitudes of the centroids of the counties along the Canadian border and measuring the distance to each area.[30] The intuition for using this variable as an IV is the following: since the Hispanic population in the U.S. tends to be concentrated towards the Mexican border, as shown in Figure 1, then as the distance from the Canadian border increases, so will the concentration of Hispanics in an area.

3.2.2 Validating the Instrumental Variable

For MinDist to be a valid IV, it must meet the exclusion restriction and the relevance condition. Importantly, we have no reason to believe that investors' tendency to invest in stocks that are local, have lottery-like characteristics, and have high past returns is systematically related to

[28] Endogeneity usually arises from three main sources: simultaneity bias, measurement error, and omitted variables. In our case, a potential causal relation is not violated by the simultaneity bias (i.e., reverse causality) since financial decisions do not determine someone's ethnicity. Similarly, measurement error should not be a primary concern because we are using census data.

[29] Since there is no overlap between high-Hispanic and high-Black areas, the minimum distance from the Canadian border is not a valid IV for predominantly high-Black areas.

[30] Figure A2 exhibits the counties along the Canadian border used to measure the IV, while Table A2 provides the names of the counties along with their FIPS code.

the IV, other than through the Hispanic investors channel. This finding suggests that our instrument satisfies the exclusion restriction.[31]

While the exclusion restriction cannot be empirically tested, we show that the instrument satisfies the relevance condition because MinDist is positive and statistically significant in the first stage regressions. The F-statistics are also high and statistically significant, suggesting that we are unlikely to have a weak instrument (Stock and Yogo, 2005).[32] However, Jiang (2017) suggests that even when an instrument passes the Stock and Yogo (2005) "weak instrument" test, an IV can lead to biased results if the instrument explains a small percentage of the endogenous variable. This is not a significant concern in our case as the first stage regressions, in Tables A5 and A6, indicate that MinDist explains a large percentage of the variation in H/W.

3.2.3 Instrumental Variable Results

Table 3 reports the IV regression results for both institutional and retail investors. The IV estimates in Panel A are in line with the OLS estimates. The coefficient in column (1) remains positive and statistically significant when we use MinDist as an instrumental variable for H/W. This suggests that Hispanic investors typically exhibit a stronger local bias. Similarly, institutional investors who live in high-Hispanic areas overweight lottery-type stocks. The last two columns show that the results for high momentum stocks, for both the past 12 and 6 months, are robust to using the instrumental variable.

The retail investor results in Panel B are consistent with the institutional investor findings. In column (1), the coefficient of H/W is large and statistically significant. When including MSA-level fixed effects, the coefficient remains statistically significant. In columns (3) and (4), we again find a strong association between a household's Hispanic culture and overweighting of lottery stocks. We also see that retail investors who reside in high H/W ZIP codes tend to

[31]The instrument satisfies the exclusion restriction as long as the only way MinDist affects households' decisions to overweight a specific subset of stocks is through the Hispanic concentration within an area. Some studies argue that the instrumental variable meets the exclusion restriction when the independent variable is regressed on the problematic regressor (H/W) and the instrumental variable (MinDist) along with other controls and the coefficient of the instrument is statistically insignificant. They suggest that the instrumental variable does not explain the dependent variable after conditioning on other independent variables. However, this regression yields biased estimates because the covariance of the problematic regressor with the error term is not equal to zero (Wooldridge, 2010; Angrist and Pischke, 2008). This is the same endogeneity issue the IV intends to address. Furthermore, since the instrumental variable is correlated with the problematic regressor (the variables need to be correlated in order to meet the relevance condition), then the coefficient on the instrumental variable is also biased (Wooldridge, 2010; Angrist and Pischke, 2008).

[32]The first stage regressions and the F-statistics can be found in Tables A5 and A6.

overweight stocks which have performed well in the past 12 and 6 months.

The IV regressions estimate a local average treatment effect rather than the population average treatment effect. In this case, our IV tests focus on institutional investors and households who live in high-Hispanic areas and invest in the stock market. The coefficients and t-statistics from the IV regressions both increased, suggesting that there is some endogeneity that biases the OLS estimates downwards.

To summarize the results from Tables 2 and 3, we observe a strong association between the Hispanic culture and investment in local, lottery, and high-momentum stocks. These findings are consistent with culture affecting investment preferences. The local bias results support our conjecture that due to their lower trust levels, Hispanics tend to favor local stocks. The lottery stock and return chasing results are consistent with "Keeping-up-with-the-Joneses" type preferences (Soriano, 1995; Hong, Jiang, Wang, and Zhao, 2014). Finally, these findings are robust to focusing on variation in Hispanic population concentration that is instrumented by each area's minimum distance to the Canadian border.

3.2.4 Controlling for Heterogeneity in Financial Sophistication

Beyond our IV approach aimed at broadly addressing concerns about omitted variables, we also directly consider whether our investor preference results are driven by financial sophistication, since less sophisticated investors could hold more concentrated portfolios (Grinblatt and Keloharju, 2001). Since retail investors are typically less financially sophisticated and tend to commit a greater number of financial mistakes, we control for this potential alternative explanation in our retail brokerage subsample.[33] Specifically, we follow the method of Rajan and Zingales (1998) and interact our Hispanic concentration and IV variables with two measures of household i's portfolio concentration, which are designed to control for varying levels of financial sophistication.

The first measure, CONC, is defined as CONC = 10 − NSTOCKS, where NSTOCKS is the number of stocks in household i's portfolio. The second measure, HERF, is a Herfindahl Index for the portfolios of all long-only investors. Ultimately, if the Hispanic concentration

[33] In our institutional investor baseline regressions, we include HERF as a control. Hence, it is unlikely that the level of an institutional investor's financial sophistication is able to explain her/his preferences for local, lottery, and high-momentum stocks.

(H/W) predicts local bias, excess weighting in lottery stocks, and return chasing, then for two investors with the same level of portfolio concentration living in different ZIP codes, we expect the investor living in the ZIP code with a higher Hispanic concentration to exhibit a stronger local bias and a higher investment in lottery and high-momentum stocks.

We report the estimates from OLS Rajan and Zingales (1998) regressions in Panel A of Table 4.[34] As expected, the interaction terms H/W×CONC and H/W×HERF are positive and statistically significant, suggesting that the Hispanic culture affects excess weighting in local, lottery, and high-momentum companies. Column (1) shows that the results for local stocks are strong. The coefficient of 60.52 is highly economically significant.[35] For example, consider two households, both with a portfolio concentration equal to the median of 7, but one lives in a ZIP code at the 75$^{\text{th}}$ percentile of Hispanic concentration (H/W = 0.102) and the other lives in a ZIP code at the 25$^{\text{th}}$ percentile of Hispanic concentration (H/W = 0.017). The household that lives in the 75$^{\text{th}}$ percentile ZIP code is predicted to invest $60.518 \times (0.102 - 0.017) \times 7 = 36.01\%$ of market weight more of its portfolio in local stocks than the household that lives in the 25$^{\text{th}}$ percentile ZIP code.

The results for the interaction of Hispanic concentration and the Herfindahl Index are similar.[36] For instance, if we consider two households with portfolios at the median level of concentration (0.536) as measured by the Herfindahl Index, the household that lives in the 75$^{\text{th}}$ percentile ZIP code is predicted to invest $659.226 \times (0.102 - 0.017) \times 0.536 = 30.03\%$ of market weight more of its portfolio in local stocks than the household that lives in the 25$^{\text{th}}$ percentile ZIP code.

Moreover, in column (3), our measure of excess weight on lottery stocks is regressed on the interaction term H/W×CONC and a vector of additional controls measured at the household and ZIP code level. We see that the coefficient of 9.73 (t-statistic = 4.51) is positive and statistically significant. The observed coefficient values are also economically significant. For example, consider two retail investors with portfolio concentrations at the median, that is CONC = 7 (HERF = 0.536). One of the retail investors lives in a ZIP code at the 75$^{\text{th}}$ percentile of

[34]The univariate regressions can be found in Table A4. The coefficients H/W×CONC and H/W×HERF are positive and statistically significant, consistent with the full specification models.

[35]Standard errors are clustered at the ZIP code-year level.

[36]The large coefficient is due to the fact that the interaction term is not standardized. The scale of H/W×HERF is much smaller than that of H/W×CONC. Looking at the net effect of living in a higher concentration ZIP code, we get similar results for both interaction terms.

Hispanic concentration and the other lives in a ZIP code at the 25$^{\text{th}}$ percentile of Hispanic concentration. The household that lives in the high H/W concentration ZIP code will invest $9.733 \times (0.102 - 0.017) \times 7 = 5.80\%$ (8.03% as measured by the Herfindahl Index) more, relative to the market weight, in lottery stocks than compared to the household which lives in the low concentration ZIP code.

Columns 5 through 10 of Table 4 present regressions of the excess weight in high-momentum stocks on our interaction terms. The H/W×CONC and H/W×HERF coefficients when the dependent variable is the excess weight on stocks in the top decile of the most recent twelve and six month periods are positive and statistically significant. The H/W×CONC for stocks that have been performing well in the past twelve months is 1.06 with a t-statistic of 2.22. For the six-month return regression, the coefficient of H/W×CONC is 2.03 with a t-statistic of 3.89. These results show that the return chasing behavior observed in Table 2 is not due to an omitted variable, such as financial sophistication, that correlates with our H/W measure.

Table 4 Panel B reports the results for the IV regressions. Column (1) suggests that the local bias results are robust to using the minimum distance from the Canadian border as an IV. Similarly, column (3) shows that when we regress our measure of excess weight on lottery stocks on H/W IV×CONC, the coefficient is positive and statistically significant. Columns (2) and (4) show that when we use HERF as a measure of portfolio concentration instead of CONC, the effect remains the same: retail investors who reside in ZIP codes with a high-Hispanic concentration tend to invest in local firms and lottery stocks.

Columns (5) to (10) report the IV regression estimates for stocks that have been performing well recently. Like the OLS regressions, the coefficients for H/W IV×CONC and H/W IV×HERF are statistically significant when the dependent variable is the excess weight on stocks in the top decile of the most recent twelve-month returns. Similarly, both coefficients are positive and statistically significant when the dependent variable is the excess weight on stocks in the top decile of the most recent six-month returns.

These results provide additional evidence that our main results are not likely to be driven by an omitted variable. Not only do we include MSA-level fixed effects that take into account time-invariant omitted variables, but we also use the minimum distance from the Canadian border to each ZIP code in our sample as an instrumental variable. Overall, we are able to establish a tight empirical relation between the Hispanic culture and relative overweighting in local, lottery, and high-momentum stocks.

3.3 Investor Herding in High-Hispanic Areas

The propensity of investors in high-Hispanic areas to invest in local, lottery, and high-momentum firms suggests the possibility that their portfolio decisions could affect stock returns. Given the "group-oriented" behavior of Hispanics, it is possible that they systematically trade in a particular set of stocks. To test whether investors located in high-Hispanic areas are more likely to herd, we use data from ANcerno Ltd. and the discount brokerage house to compute the following herding measure (Lakonishok, Shleifer, and Vishny, 1992; Barber, Odean, and Zhu, 2009b):

$$HM_{i,t} = |p_{i,t} - E[p_{i,t}]| - E|p_{i,t} - E[p_{i,t}]| \tag{1}$$

for stock i in month t. $|p_{i,t} - E[p_{i,t}]|$ is the magnitude of the proportion of purchases to the total number of trades in stock i during month t minus the proportion of all purchases during month t.

The first term of the herding measure captures the degree to which the decisions of different retail investors to buy or sell a stock are correlated. The latter term of the measure, $E|p_{i,t} - E[p_{i,t}]|$, is used to adjust for the fact that more variation in the proportion of buys is expected in stocks that have fewer trades.[37] We compute this measure for all stocks in the brokerage data set that have at least 10 trades in month t. Then, during each month, we average the herding measure across all stocks. Motivated by our stock preference results, we test, within each Hispanic concentration quintile, whether investors herd in local, lottery, and high-momentum stocks. We expect to find a higher buying pressure in these types of stocks in the quintile with the highest concentration of Hispanics.

Table 5 provides evidence that investors in high-Hispanic areas herd within local, lottery, and high-momentum firm categories. The estimates in Panel A indicate that institutional investors in both high- and low-Hispanic areas herd in local firms, as the herding measures are positive and statistically significant. Importantly, the herding estimate for the high-Hispanic quintile, 0.246 (p-value = 0.00), is higher than the herding estimate for the low-Hispanic quintile. Similarly,

[37]See Lakonishok, Shleifer, and Vishny (1992) for further details about this measure.

the systematic buying pressure for lottery stocks is higher for institutional investors located in high-Hispanic areas. In particular, the coefficient for lottery stocks in high-Hispanic areas is 0.250 (p-value = 0.00), while the coefficient estimate in low-Hispanic areas is 0.234 (p-value = 0.00).

In the high-Hispanic quintile, the herding measures for stocks with high-momentum returns during the past 12 or 6 months are 0.240 (p-value = 0.00) and 0.227 (p-value = 0.00), respectively. These herding measures are larger than the corresponding herding measures in the low-Hispanic quintile, 0.222 (p-value = 0.00) and 0.218 (p-value = 0.00), respectively.

The results in Panel B are consistent with these findings. They show that on average, retail investors in high-Hispanic areas herd in local firms. The herding measure is positive and statistically significant (= 0.058, p-value = 0.00); but it is negative and statistically insignificant in low-Hispanic areas (-0.022, p-value = 0.88). The evidence also indicates that investors in high-Hispanic areas are potentially slightly more likely to herd in lottery stocks (herding estimate = 0.017, p-value = 0.04) than investors in low-Hispanic areas (herding estimate = 0.016, p-value = 0.07), although this difference is not statistically significant.[38]

Additionally, the estimates suggest that retail investors systematically buy high-momentum firms. In the high-Hispanic quintile, the herding measures for stocks that have experienced high-momentum returns in the past 12 and 6 month periods are consistently positive and statistically significant: 0.039 (p-value = 0.00) and 0.041 (p-value = 0.00), respectively. More importantly, these coefficients are larger than the respective herding measures of high-momentum stocks in low-Hispanic areas. For instance, the herding measure for high-momentum stocks in the past 12 months is 0.017 (p-value = 0.03) and for the past 6 months is 0.023 (p-value = 0.01).

Collectively, the micro-level evidence suggests that institutional and retail investors in high-Hispanic areas have a higher propensity to herd in local, lottery, and high-momentum stocks.[39] Since the trading decisions of investors in these types of stocks appear to be systematic, investors in high-Hispanic areas *may* affect the returns of these stocks. We test this conjecture in the

[38]As shown in the last column of Table 5, not all differences are statistically significant, but the overall pattern is consistent with our broad conjectures.

[39]It is difficult to make direct comparisons of retail and institutional herding estimates across samples. Nevertheless, the magnitudes of our herding estimates for the retail investor sample are comparable to those reported in Barber, Odean, and Zhu (2009b) and Barber, Odean, and Zhu (2009a). Also, as Barber, Odean, and Zhu (2009a) note, the herding measures using monthly data are larger than those reported in the literature using quarterly institutional holdings data.

next section.

3.4 Return Comovement in High-Hispanic Areas

The systematic purchases of local, lottery, and high-momentum firms may induce a common factor in the returns of these securities that cannot be explained by fundamentals (Barberis, Shleifer, and Wurgler, 2005). In addition, Eun, Wang, and Xiao (2015) suggest that in countries with culturally tight and "group-oriented" traits, stock returns experience higher comovement. Therefore, we expect the comovement of returns to be higher in high-Hispanic areas.

To test whether stock returns comove more in high-Hispanic areas, we follow Pirinsky and Wang (2006) and construct a habitat-portfolio index, which consists of a value-weighted portfolio of firms that are local to institutional investors, lottery-type stocks, and high-momentum companies for each Hispanic quintile.[40] We then estimate the following stock-level time-series regression:

$$\begin{aligned} r_{i,t} = \alpha_i + \beta_1 r_{t,Habitat_{-i}} + \beta_2 r_{t,Mkt_{-Habitat,-i}} + \beta_3 r_{t,SMB} \\ + \beta_4 r_{t,HML} + \beta_5 r_{t,MOM} + \beta_6 r_{t,Liquidity} + \epsilon_{i,t}, \end{aligned} \tag{2}$$

where $r_{i,t}$ is the monthly excess return of a particular stock and $r_{t,Habitat_{-i}}$ is the monthly return of the value-weighted habitat portfolio. $r_{t,Mkt_{-Habitat,-i}}$, $r_{t,HML}$, and $r_{t,MOM}$ are the three Fama-French factors. We also modify this specification to account for short- and long-term reversals (De Bondt and Thaler, 1985; Jegadeesh, 1990; Conrad and Kaul, 1998):

$$\begin{aligned} r_{i,t} = \alpha_i + \beta_1 r_{t,Habitat_{-i}} + \beta_2 r_{t,Mkt_{-Habitat,-i}} + \beta_3 r_{t,SMB} \\ + \beta_4 r_{t,HML} + \beta_5 r_{t,STR} + \beta_6 r_{t,LTR} + \epsilon_{i,t}. \end{aligned} \tag{3}$$

To avoid multicollinearity, we exclude firm i's return from the habitat and market portfolios and the firms included in the habitat portfolio from the market portfolio. The coefficient of interest is β_1, the beta of the habitat portfolio. Since investors are more likely to herd in local, lottery, and high-momentum firms in high-Hispanic areas, we expect the habitat beta for these types of stocks to be highest in the 5$^{\text{th}}$ H/W quintile.

The estimates for equation 2 in Panel A show that the habitat beta of firms in high-Hispanic areas that are local to institutional investors is 0.576 (t-statistic = 10.14), while the habitat beta

[40] The results are robust to and become stronger using an equally-weighted portfolio.

for low-Hispanic areas is 0.120 (t-statistic = 2.08). This suggests that the comovement of local stocks in the 5th H/W quintile is greater than the comovement in the 1st quintile. Consistent with this notion, the difference of 0.456 is positive and statistically significant (p-value = 0.00).[41]

For lottery stocks, the habitat beta in the 5th quintile is 0.838 (t-statistic = 13.45), and the level of comovement decreases monotonically across the Hispanic quintiles. The lowest habitat beta for lottery stocks is 0.321 (t-statistic = 5.27) and occurs in the 1st quintile. The difference between the highest and lowest habitat beta is 0.517 (p-value = 0.00).

We find similar results for stocks that have experienced high-momentum returns in the past 12 and 6 month periods: the habitat betas for the 5th H/W quintile are greater than the habitat betas for the 1st quintile. For example, the habitat betas for the 12 and 6 month momentum returns for the high-Hispanic quintile are 0.467 (t-statistic = 11.83) and 0.437 (t-statistic = 14.04), respectively. In the low-Hispanic quintile, the habitat beta for the 12 month momentum return is statistically insignificant and for the 6 month momentum return is 0.078 (t-statistic = 1.90).

As additional evidence that the comovement is likely to be driven by Hispanics' stock preferences, we create an index within each quintile that measures how well each stock meets these characteristics: being local, having a high degree of lottery-like characteristics, and experiencing high-momentum returns. We then divide stocks into quintiles using this index.[42] Stocks that are local to institutional investors, have lottery-like characteristics, and experience high 12 month momentum returns are classified into the "High Index" quintile. Conversely, stocks that are not nearby institutional investors, do not have lottery-like characteristics, and do not have high returns over the previous 12 months are classified into the "Low Index" quintile.

The estimates in Table 6, Panel A show that within the high-Hispanic quintile, the comovement along the index quintiles decreases monotonically with the index. The habitat beta for the "High Index" quintile is 0.501 (t-statistic = 9.78), while the habitat beta for the "Low Index" quintile is 0.224 (t-statistic = 4.00). The difference between these two betas is positive and sta-

[41] The results are robust to using firms that are local to retail investors, rather than institutional investors. The main findings also remain unchanged if we use average institutional or retail ownership.

[42] Specifically, each month we sort stocks into vigintiles (20 bins) by their price, idiosyncratic volatility, idiosyncratic skewness, and 12 month momentum return. For each stock, the vigintile bin scores are added and divided by 4 to produce a score between 1 and 20. For local stocks, we use a binary variable to determine whether a firm is local. The results are robust to using the average institutional ownership or retail ownership. For more details on this measure, please refer to Kumar, Page, and Spalt (2016).

tistically significant, 0.277 (p-value = 0.00). In the low-Hispanic quintile, the difference between the two habitat betas is statistically insignificant. If we compare the habitat betas for the high and the low-Hispanic areas, we see that the difference between these two is always positive and statistically significant, confirming that firms in high-Hispanic areas tend to experience higher comovement of returns.

Panel B exhibits the results for equation 3, which includes the three Fama-French factors along with the short- and long-term reversal factors. The habitat betas for the 5^{th} Hispanic concentration quintile are systematically larger than those for the 1^{st} Hispanic concentration quintile and these differences are statistically significant. These findings suggest that local, lottery, and high-momentum stocks headquartered in high-Hispanic areas experience the strongest return comovement.

The index quintiles within the high-Hispanic areas show that as stocks become more non-local, less lottery-like, and have lower momentum returns, the comovement of returns also decreases. The difference between the "High Index" and the "Low Index" habitat beta is positive and statistically significant, 0.271 (p-value = 0.00). For the low-Hispanic quintile, there is no clear pattern for the comovement of returns and the difference between the "High Index" and the "Low Index" habitat beta is statistically insignificant. Overall, these results suggest that the culturally-driven systematic trades of investors in high-Hispanic areas can result in excess comovement of returns that cannot be explained by fundamentals.

4 Hispanic Culture and Cross-Sectional Returns

In this section, we study the cross-sectional asset pricing effects associated with Hispanic investors' stock preferences and correlated trading. Specifically, we analyze whether their relatively high demand for local assets affects local stock prices, whether the negative lottery-stock premium is more negative in high-Hispanic areas, and if Hispanics' return chasing behavior affects the returns of local momentum portfolios.

For the asset pricing tests, we use data from CRSP and Compustat from January 1970 to December 2018. However, since the 13(f) Thomson Reuters database is only available starting in 1980, the analysis for local stocks is performed from January 1980 to December 2018.

4.1 Hispanic Concentration and Local Stock Returns

The institutional and retail investor preference findings suggest that Hispanic investors have a stronger local bias. Therefore, it is possible that in areas where there is a high concentration of Hispanic residents, investors who overweight local stocks drive up the prices of local firms and generate higher average short-term returns (Hong, Kubik, and Stein, 2008).

To test this conjecture, we use a bivariate dependent sort, where we first sort firms into quintiles based on their local, MSA-level H/W measure.[43] Then, within each quintile, we sort firms into quintiles based on their previous quarter change in local institutional investor holdings (e.g., investors who are located less than 60 miles away from the company's headquarter), "ΔIO."[44]

Table 7, Panel A presents the value-weighted excess returns (i.e., in excess of the risk-free rate) for each quintile. As hypothesized, the results suggest that within the high-Hispanic quintile, "High ΔIO" stocks earn higher returns than "Low ΔIO" stocks; the difference of 0.80% is statistically significant with a t-statistic of 2.56. Conversely, within the low-Hispanic quintile, the difference between "High ΔIO" and "Low ΔIO" firms is negative and statistically insignificant. Further, the H-L row suggests that the difference in returns in high-Hispanic areas is greater than the difference in returns in low-Hispanic areas, 0.86% (t-statistic of 2.54).

In Panel B, we create a zero-cost portfolio for each Hispanic quintile. It consists of going long on firms in the fifth quintile and short on firms in the first quintile based on previous quarter changes in local institutional ownership. All portfolios are value-weighted. We control for known risk factors by regressing the portfolio excess returns on the three Fama and French (1993) factors, the short-term reversal factor (De Bondt and Thaler, 1985; Jegadeesh, 1990), and the long-term reversal factor (De Bondt and Thaler, 1985; Jegadeesh, 1990; Conrad and Kaul, 1998).

In the high-Hispanic quintile, the estimates suggest that the long portfolio has a positive and statistically significant alpha of 0.43% (t-statistic = 1.73), while the short portfolio has a negative and statistically alpha of -0.52% (t-statistic = -2.55). As a result, the Long-Short

[43] We aggregate ZIP code-level population statistics to the MSA-level using the decennial census data set. For the decade beginning in 1970, census data is only available at the county-level, so we use data at this geographical level.

[44] The sample is restricted to firms that have at least one institutional investor located within a 60 mile radius of their headquarter (e.g., firms that have positive local institutional ownership values).

strategy generates a positive and statistically significant abnormal return of 0.95% (t-statistic = 2.86). We also find that the alpha of the high-Hispanic quintile is higher than the risk-adjusted return for the low-Hispanic quintile, which is statistically insignificant. The "Alpha Difference" row suggests that these two abnormal returns are statistically different, confirming that the alpha is greater in the 5$^{\text{th}}$ Hispanic quintile.

Figure 2 plots the time series evolution of the differences in alphas between the zero-cost portfolios in the high- and low-Hispanic quintiles. It shows that the difference in abnormal returns is no longer statistically significant after 8 months. Collectively, the evidence in this section provides support for our conjecture. Since investors in high-Hispanic areas have a stronger local bias, then their increased demand for local assets generates higher short-term average returns.

4.2 Lottery-Stock Premium

It is possible that Hispanic investors' preferences affect the stock returns of lottery-type stocks. In particular, the magnitude of the lottery-stock premium could be more negative for firms located in high-Hispanic areas.

We analyze the returns of lottery and non-lottery stocks in high- and low-Hispanic areas in Table 8. Specifically, we create a bivariate dependent sort, where we first classify firms into lottery and non-lottery stocks. We then sort the companies within each group into quintiles based on their local Hispanic concentration measure. Panel A presents the value-weighted excess returns (i.e., in excess of the risk-free rate) for each classification. As conjectured, the results suggest that the lottery-stock premium is more negative when there is a higher concentration of Hispanics in an area. The difference in the average returns between the high- and the low-Hispanic areas is negative and statistically significant, -0.55 (t-statistic = -1.94). Conversely, there is a positive association between the returns of non-lottery stocks and the local concentration of Hispanic residents. The difference between the high- and low-Hispanic areas is not statistically significant.

We also test the difference in the average returns between lottery and non-lottery stocks in high and low-Hispanic areas. We find that on average, lottery-type stocks earn lower returns. Importantly, the magnitude of the difference is higher for companies headquartered in high-

Hispanic areas. This evidence is consistent with the literature, which suggests that lottery stocks typically have a negative premium (Barberis and Huang, 2008; Kumar, Page, and Spalt, 2011).

We also examine if these patterns can be explained by known risk factors. In Panel B, we create a trading strategy where each quintile consists of a value-weighted portfolio. The zero-cost portfolio consists of going long on the high-Hispanic portfolio and short on the low-Hispanic portfolio. We control for several risk factors, which include the Fama and French (1993) three factors, the short-term factor (De Bondt and Thaler, 1985; Jegadeesh, 1990), and the long-term reversal factor (De Bondt and Thaler, 1985; Jegadeesh, 1990; Conrad and Kaul, 1998). We expect for the trading strategy to yield a negative abnormal return when the sample is restricted to lottery stocks.

As conjectured, the estimates suggest that the trading strategy for lottery-type stocks generates a negative and statistically significant alpha of -0.53 (t-statistic $= -2.17$) per month. It is important to highlight that this is not a trend that is general to all stocks, as the alpha for non-lottery stocks is positive and statistically insignificant. Additionally, we test whether the differences in alphas between lottery and non-lottery stocks are statistically significant. The results show that the alphas of lottery stocks tend to be lower. The "Alpha Difference" between high- and low-Hispanic areas is also negative and statistically significant, -0.68 (p-value $= 0.02$), indicating that the magnitude of the negative lottery-stock premium is more negative in high-Hispanic areas.

4.3 Momentum Returns in High-Hispanic Areas

Building upon the investor preference results, we now test whether momentum returns are higher in high-Hispanic areas. We analyze both raw and risk-adjusted momentum returns.

4.3.1 Raw Momentum Returns

The propensity for investors to herd and chase returns is one mechanism that can yield the momentum effect. By purchasing stocks that have recently done well, investors may drive prices up even further. If at least some portion of these purchases are financed by selling stocks that have recently done poorly, the momentum effect may be amplified. Our earlier results showing that investors in high-Hispanic concentration ZIP areas herd in stocks with high past returns suggest that the momentum effect may be stronger for firms headquartered in areas with a high-Hispanic concentration.

We create momentum portfolios using a bivariate dependent sort. First, we sort stocks into quintiles based on the most recent observation of our Hispanic concentration measure. Then, at the beginning of each month, we sort firms in each of these quintiles into deciles based on their previous six-month returns. Following Jegadeesh and Titman (1993), we create equally weighted portfolios holding the stock for the subsequent six-months, skipping a month between formation and holding periods to avoid microstructure biases (i.e., short-term reversals, bid-ask bounce, lead-lag reaction effects and price pressure). The "winners" portfolio is the equal weighted portfolio of firms in the top decile of the return distribution while the "losers" portfolio contains firms in the bottom decile. Momentum portfolios are created by buying winners and selling losers.

Table 9, Panel A presents the results from this bivariate dependent sort. As shown in the first column, the returns to the winners are unrelated to the H/W ratio. Nonetheless, returns to losers decrease almost monotonically as the Hispanic concentration increases, with the exception of the third quintile. Consequently, the high-Hispanic concentration momentum portfolio (MOM_5) is greater than the low concentration momentum portfolio (MOM_1). The difference ($\text{MOM}_5 - \text{MOM}_1$) is statistically significant with a t-statistic of 1.74 (p-value = 0.08).

Figure 3 provides a visual representation of the findings from Table 9, Panel A. The figure displays the cumulative monthly log-returns for winners and losers in Hispanic concentration quintiles 1 and 5, as well as the market and the risk-free returns as benchmarks. The "winners" portfolios in both quintiles have similar returns. For instance, a dollar invested in winners headquartered in areas with a low-Hispanic concentration at the beginning of 1970 is worth \$151.05 in 2018. Likewise, a dollar invested in winners headquartered in areas with a high-Hispanic concentration is worth \$137.63.

The majority of the difference in performance comes from the "losers" portfolios. The final dollar value from holding losers in the highest quintile of concentration is $0.83/25 = 3.32\%$ of the dollar value of holding losers in the bottom quintile of H/W. These results are in line with the findings of Moskowitz and Grinblatt (1999), who show that momentum profits are due to the short leg of the strategy. In addition, the total return on the market is \$125.46, lower than both "winners" portfolios.

Overall, the findings in this section support our conjecture. We find evidence in line with the preferences of local investors generating geographic segmentation in momentum returns. More specifically, we find that momentum returns are concentrated (nonexistent) in areas of high (low) Hispanic concentration.

4.3.2 Momentum Profits using Various Factor Models

Since the results from Table 9, Panel A show unconditional means for the momentum strategies, in this section we investigate whether these results can be explained by known risk factors. Table 9, Panel B presents risk-adjusted estimates of momentum returns for firms located in areas of low and high-Hispanic concentration. Momentum returns for quintiles 1 and 5 are regressed on the three Fama-French factors (Fama and French, 1992), the short-term reversal factor (Jegadeesh, 1990; Conrad and Kaul, 1998) and the long-term reversal factor (De Bondt and Thaler, 1985; Jegadeesh, 1990; Conrad and Kaul, 1998).

The results show that the momentum profits are robust: neither the CAPM, the three Fama-French factors, nor the short- or long-term reversal factors can explain the abnormal returns across H/W quintiles. The alphas of the momentum portfolios of firms located in high-Hispanic concentration areas are positive and statistically significant in all of the models. They range from 0.85% per month in the CAPM model to 1.13% per month in the three Fama-French plus reversal factors model. The differences in alphas within each model across high and low quintiles are also large. On average, risk-adjusted momentum returns from the high-Hispanic concentration areas tend to be twice as large as the momentum returns from the low-Hispanic concentration areas. These differences in alphas are consistently positive and statistically significant at the 1% level. In addition, when the short- and long-term reversal factors are included in the specifications, they improve the fit of the model, suggesting that there could be a cultural or geographic component to these factors' ability to explain momentum profits.

4.4 Additional Evidence and Robustness Tests

Our results so far indicate that the Hispanic culture influences investors' stock preferences, and consequently, asset prices in high-Hispanic areas. In the Internet Appendix, we conduct several additional tests to further support our hypotheses. In particular, we extend our analysis to examine the effect of Hispanic culture on real estate returns and document that prices in high-Hispanic areas exhibit higher price run-ups and subsequent downturns. We also test an alternative distance measure and find that our IV results hold when using the minimum distance from the Mexican border as an alternative instrument. Finally, we show that our results are robust to controlling for additional measures of financial sophistication as well as an alternative measure for an area's Hispanic concentration.

5 Summary and Conclusions

This paper examines the impact of culture on portfolio decisions and aggregate financial market outcomes. Specifically, we study the effects of the Hispanic culture on investors' stock preferences and, in turn, whether these systematic choices of investors affect asset prices. We focus on the Hispanic culture because it is the largest and fastest growing community in the U.S. In addition, it has several unique features and consumption patterns that aid in identifying the impact of culture on financial markets better than through other ethnicities.

We find that both institutional and retail investors living in areas with a high concentration of Hispanic residents exhibit a strong preference for local, lottery, and high-momentum stocks. Importantly, through our instrumental variable (IV) analysis, we show that these effects cannot be attributed to unobserved heterogeneity correlated with Hispanic population concentration. We also find that Hispanic investors' preferences affect stock prices and returns. More specifically, our evidence suggests that investors in high-Hispanic areas are more likely to herd in local, lottery, and high-momentum stocks, generating excess comovement in returns that cannot be explained by firm fundamentals. These investors' preferences also drive the short-term returns of local stocks and amplify the magnitude of both the lottery stock premium and momentum returns among firms located in high-Hispanic areas.

Collectively, these results provide a link between the culture of local residents, their portfolio choices, and stock return patterns. In future work, it would be interesting to examine the extent to which the impact of culture extends beyond institutional and retail investors. For example, the forecasting behavior of sell-side analysts may be influenced by the local culture where analysts of Hispanic origin may issue forecasts consistent with their cultural background. Further, analysts with a more "group-oriented" Hispanic culture may exhibit a lower propensity to issue bold forecasts because they do not want to deviate from the consensus. If the local culture affects analysts and other groups of market participants in a significant manner, cultural factors may have broader impact on aggregate economic outcomes. We examine these and related issues in our ongoing research.

References

Abel, Andrew B, 2003, The effects of a baby boom on stock prices and capital accumulation in the presence of social security, *Econometrica* 71, 551–578.

Ang, Andrew, and Angela Maddaloni, 2005, Do demographic changes affect risk premiums? Evidence from international data, *The Journal of Business* 78, 341–380.

Angrist, Joshua D, and Jörn-Steffen Pischke, 2008, *Mostly harmless econometrics: An empiricist's companion* (Princeton University Press).

Antoniou, Constantinos, John A Doukas, and Avanidhar Subrahmanyam, 2010, Sentiment and momentum, Working Paper.

Avramov, Doron, Si Cheng, and Allaudeen Hameed, 2015, Time-varying liquidity and momentum profits, *Journal of Financial and Quantitative Analysis.*

Bakshi, Gurdip S, and Zhiwu Chen, 1994, Baby boom, population aging, and capital markets, *Journal of Business* pp. 165–202.

Barber, Brad M, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *The Journal of Finance* 55, 773–806.

——— , and Ning Zhu, 2009a, Do retail trades move markets?, *The Review of Financial Studies* 22, 151–186.

——— , 2009b, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066–2100.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Becker, Gary S, 1996, *Accounting for Tastes* (Harvard University Press).

Becker, Gary Stanley, Kevin M Murphy, Iván Werning, et al., 2000, *Status, lotteries and inequality* (George J. Stigler Center for the Study of the Economy and the State, University of Chicago).

Brenner, Reuven, 1983, *History–The human gamble* (Chicago: U. of Chicago Press).

Brunk, Gregory G, 1981, A test of the Friedman-Savage gambling model, *The Quarterly Journal of Economics* 96, 341–348.

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Cartagena, Chiqui, 2013, *Latino Boom II: Catch the biggest demographic wave since the Baby Boom* (Worthy Shorts, Incorporated).

Cheon, Yong-Ho, and Kuan-Hui Lee, 2017, Maxing out globally: Individualism, investor attention, and the cross section of expected stock returns, *Management Science*.

Chui, Andy C.W., Sheridan Titman, and K.C. John Wei, 2010, Individualism and momentum around the world, *The Journal of Finance* 65, 361–392.

Clotfelter, Charles T, and Philip J Cook, 1990, On the economics of state lotteries, *Journal of Economic Perspectives* 4, 105–119.

Conrad, Jennifer, and Gautam Kaul, 1998, An anatomy of trading strategies, *Review of Financial Studies* 11, 489–519.

Cooper, Michael J, Roberto C Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *The Journal of Finance* 59, 1345–1365.

Coval, Joshua D, and Tobias J Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

——— , 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact?, *The Journal of Finance* 40, 793–805.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *The Journal of Finance* 45, 379–395.

DellaVigna, Stefano, and Joshua M Pollet, 2007, Demographics and industry returns, *American Economic Review* 97, 1667–1702.

Demarzo, Peter M., Ron Kaniel, and Ilan Kremer, 2004, Diversification as a public good: Community effects in portfolio choice, *The Journal of Finance* 59, 1677–1716.

Eun, Cheol S, Lingling Wang, and Steven C Xiao, 2015, Culture and R2, *Journal of Financial Economics* 115, 283–303.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Fconomics* 33, 3–56.

Fernandez, Raquel, and Alessandra Fogli, 2009, Culture: An empirical investigation of beliefs, work, and fertility, *American Economic Journal: Macroeconomics* 1, 146–177.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti, 2004, Mothers and sons: Preference formation and female labor force dynamics, *Quarterly Journal of Economics* pp. 1249–1299.

Friedman, Milton, and Leonard J Savage, 1948, The utility analysis of choices involving risk, *Journal of Political Economy* 56, 279–304.

Gali, Jordi, 1994, Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices, *Journal of Money, Credit and Banking* 26, 1–8.

Goyal, Amit, 2004, Demographics, stock market flows, and stock returns, *Journal of Financial and Quantitative Analysis* 39, 115–142.

Grinblatt, Mark, and Matti Keloharju, 2001, How distance, language, and culture influence stockholdings and trades, *The Journal of Finance* 56, 1053–1073.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2003, People's opium? Religion and economic attitudes, *Journal of Monetary Economics* 50, 225–282.

——— , 2004a, Cultural biases in economic exchange, Discussion paper, National Bureau of Economic Research.

———, 2004b, The role of social capital in financial development, *American Economic Review* 94, 526–556.

———, 2006, Does culture affect economic outcomes?, *Journal of Economic Perspectives* 20, 23–48.

———, 2008, Trusting the stock market, *The Journal of Finance* 63, 2557–2600.

Hilary, Gilles, and Kai Wai Hui, 2009, Does religion matter in corporate decision making in america?, *Journal of Financial Economics* 93, 455–473.

Hispanic Wealth Project, 2019, State of hispanic wealth report 2019, Available at: `http://hispanicwealthproject.org/downloads/2019-SHWR-Annual-Report.pdf`.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2004, Social interaction and stock-market participation, *The Journal of finance* 59, 137–163.

———, 2008, The only game in town: Stock-price consequences of local bias, *Journal of Financial Economics* 90, 20–37.

Hong, Harrison G., Wenxi Jiang, Na Wang, and Bin Zhao, 2014, Trading for status, *Review of Financial Studies*.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *The Journal of Finance* 45, 881–898.

———, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* 48, 65–91.

Jiang, Wei, 2017, Have instrumental variables brought us closer to the truth, *The Review of Corporate Finance Studies* 6, 127–140.

Karolyi, G Andrew, 2016, The gravity of culture for finance, *Journal of Corporate Finance* 41, 610–625.

Korzenny, Felipe, and Betty Ann Korzenny, 2011, *Hispanic Marketing* (Routledge).

Kumar, Alok, 2009, Who gambles in the stock market?, *The Journal of Finance* 64, 1889–1933.

——— , Jeremy K. Page, and Oliver G. Spalt, 2011, Religious beliefs, gambling attitudes, and financial market outcomes, *Journal of Financial Economics* 102, 671–708.

Kumar, Alok, Jeremy K Page, and Oliver G Spalt, 2016, Gambling and comovement, *Journal of Financial and Quantitative Analysis* 51, 85–111.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Liu, Laura Xiaolei, and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Markus, Hazel R, and Shinobu Kitayama, 1991, Culture and the self: Implications for cognition, emotion, and motivation., *Psychological Review* 98, 224.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum?, *The Journal of Finance* 54, 1249–1290.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite heteroskedasticity and auto-correlation consistent variance-covariance matrix, *Econometrica* 55, 703–708.

Osili, Una Okonkwo, and Anna L Paulson, 2008, Institutions and financial development: Evidence from international migrants in the United States, *The Review of Economics and Statistics* 90, 498–517.

Pastor, Lubos, and Robert F Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Pirinsky, Christo, and Qinghai Wang, 2006, Does corporate headquarters location matter for stock returns?, *The Journal of Finance* 61, 1991–2015.

Poterba, James M, 2001, Demographic structure and asset returns, *Review of Economics and Statistics* 83, 565–584.

Rajan, Raghuram G., and Luigi Zingales, 1998, Financial dependence and growth, *American Economic Review* 88, pp. 559–586.

Ronen, Simcha, and Oded Shenkar, 2013, Mapping world cultures: Cluster formation, sources and implications, *Journal of International Business Studies* 44, 867–897.

Soriano, Marcel, 1995, Latinos in rehabilitation: Implications for culturally-appropriate counseling, *NARPPS Journal* 10, 67–72.

Stambaugh, Robert F, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Stock, James H, and Motohiro Yogo, 2005, Testing for weak instruments in linear IV regression, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* p. 80.

Stranahan, Harriet A, and Mary O Borg, 1998, Separating the decisions of lottery expenditures and participation: A truncated Tobit approach, *Public Finance Review* 26, 99–117.

Transparency International, 2007, *Global Corruption Report 2007: Corruption in Judicial Systems* . Transparency International Global Corruption Reports (Cambridge University Press).

Vincenty, Thaddeus, 1975, Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations, *Survey Review* 23, 88–93.

Weaver, Charles N, 2003, Confidence of Mexican Americans in major institutions in the United States, *Hispanic Journal of Behavioral Sciences* 25, 501–512.

Wooldridge, Jeffrey M, 2010, *Econometric analysis of cross section and panel data* (MIT press).

Figure 1
County-Level Hispanic Concentration in the U.S.

This figure plots the Hispanic concentration of a county as measured by H/W, the ratio of the Hispanic population to the White population in a county. Concentration is measured at the beginning of each decade based on the most recent census and is held constant for that decade. The decennial censuses used in this study start in 1970 and end in 2010. Only the counties where there is at least one firm headquartered are shown.

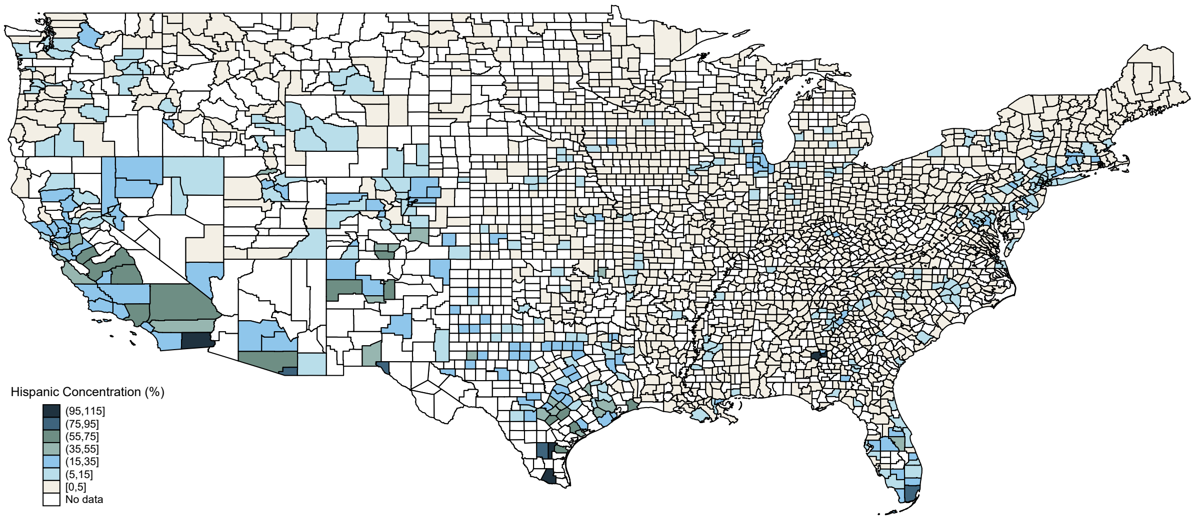

Figure 2
Time Series Performance of the Trading Strategies for Local Stocks

This figure plots the difference in alphas of the trading srategies between the high and low H/W areas in time period t+k. The trading strategies for local stocks consist of a bivariate dependent sort where we first sort stocks into quintiles based on their H/W MSA value. We then sort stocks into quintiles based on their previous quarter change in local institutional ownership (IO). The zero-cost strategy consists of going long on the top quintile based on previous quarter change in local IO and short on the bottom quintile based on previous quarter change in local IO, for each H/W percentile. We regress the portfolio returns on a set of factors, which include the market excess return (RMRF), size (SMB), value (HML), short-term reversal factor (STR) and long-term reversal factor (LTR). The p-value testing equality of alphas is presented above each bar.

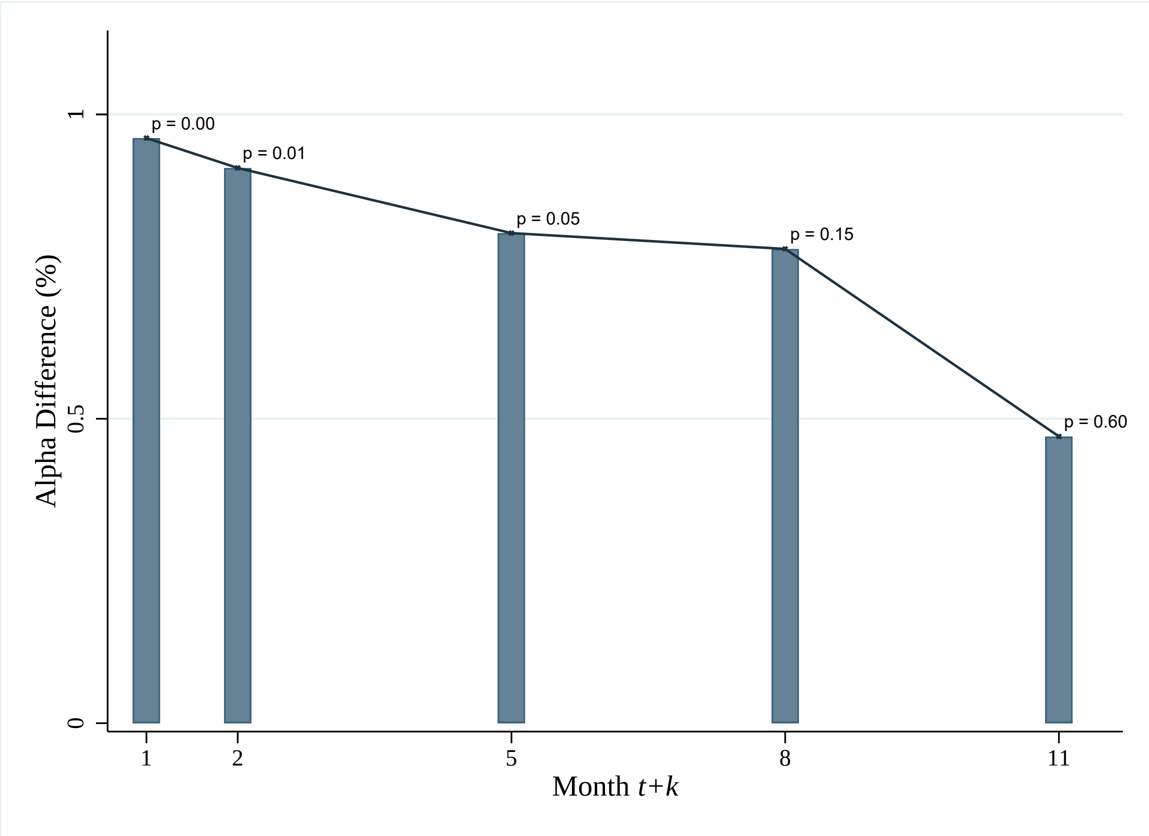

Figure 3
Cumulative Gains for Winner and Loser Portfolios

This figure plots the cumulative returns for different momentum portfolios from 1970 to 2018. Specifically, we have four momentum portfolios corresponding to stocks headquartered in either areas of high- or low-Hispanic concentration (as measured by H/W) and winners or losers. Concentration is measured at the beginning of each decade based on the most recent census and is held constant for that decade. The blue line is low-Hispanic winners and the dark red line is low-Hispanic losers. The green line is high-Hispanic winners and the yellow line is high-Hispanic losers. For comparison, the market return (light blue line) and risk free rate (1 year T-bill, red line) are also plotted.

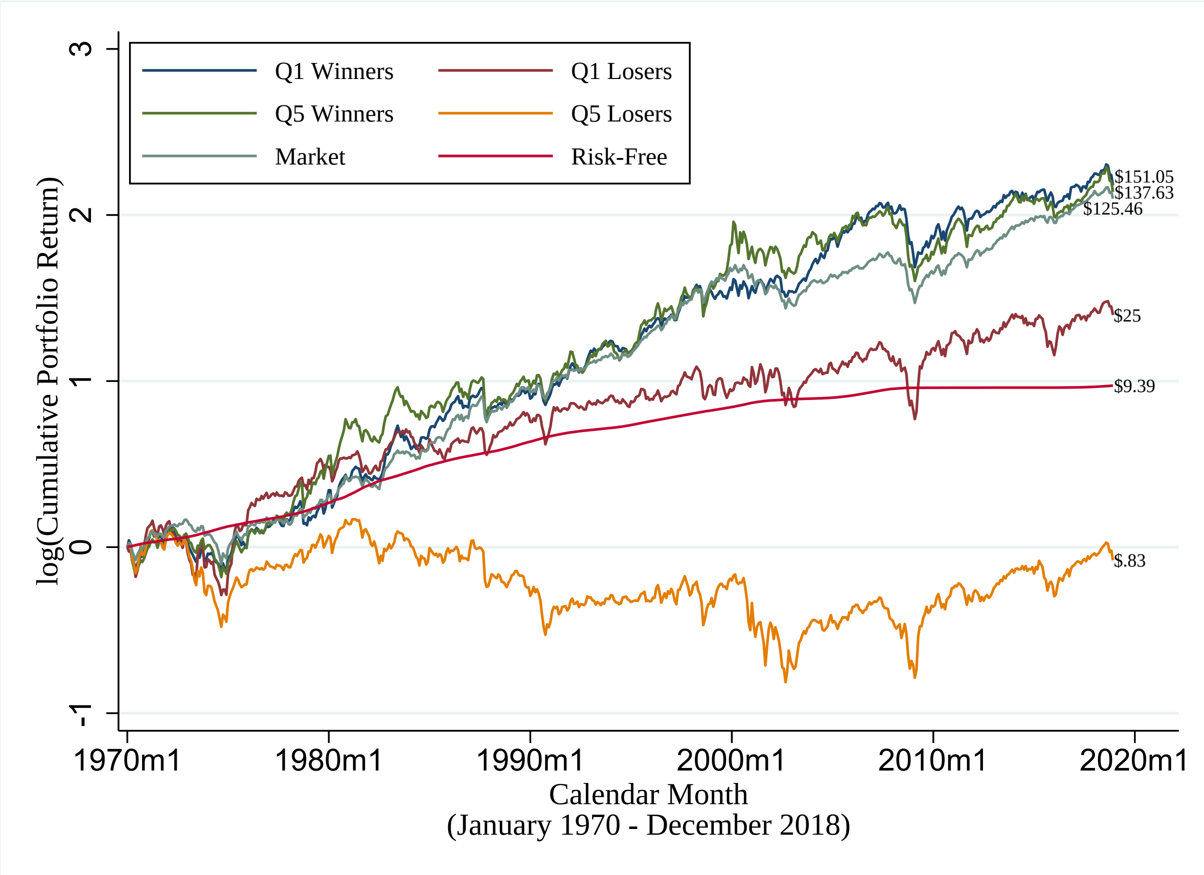

Table 1
Summary Statistics

This table presents summary statistics for the data used in the paper. Panel A presents investor-level summary statistics. The retail investor data is from a large discount brokerage house and covers the from January 1991 to November 1996. The institutional investor sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. Age is the age of the head of the household. Married is the percentage of married households. Own Home gives the percentage of investors who own their home. Male gives the percentage of households for which the head is male. Income is the household income as calculated using nine income categories with midpoints (in thousands) of 7.5, 17.5, 25, 35, 45, 62.5, 87.5, 112.5 and 250. The income of the household is assumed to be the midpoint value. Number of Securities is the number of securities owned by the household in their brokerage account. Portfolio Value is the total value of the portfolio (in thousands). Average Security Value is the average dollar value (in thousands) of all securities held in the brokerage account, while the Sharpe Ratio is the average Sharpe Ratio of the brokerage account over the six year sample period. HERF is the Herfindahl Index of the investor's portfolio. Panel B gives ZIP code-level demographic summary information for all U.S. ZIP codes, according to the 1990 census. Persons is the number of people living in each ZIP code, while Median Age, Median Income and Median Education are the median age, median household income (in thousands), and median level of education (in years) in each ZIP code. Minority is the percentage of the population classified as belonging to a minority (non-white) in each ZIP code. Male-female is the ratio of men to women in each ZIP code. Urban is the percentage of people residing in each ZIP code the Census Bureau classifies as living in an urban setting. Hispanic and Black are the percentage of people identifying as having Hispanic or Black ancestry in each ZIP code. Panel C presents the same ZIP code-level demographic summary information as presented in Panel B, but only for those ZIP codes in which at least one household from the brokerage data resides. Panel D provides summary statistics for all U.S. counties, using the decennial census from 1980, 1990, 2000, and 2010. Panel E provides similar demographic information as in Panel D, but only for those counties in which at least one institutional investor is located. The variable definitions are also available in the Appendix, Table A1.

Panel A: Investor-level portfolio characteristics

Variable	Mean	Median	Std Dev	10^{th} Pctl	25^{th} Pctl	75^{th} Pctl	90^{th} Pctl	N
Retail Investors								
Age of Head	40.9	46.0	22.9	0.0	34.0	56.0	68.0	51955
Married (%)	73.5	100.0	44.1	0.0	0.0	100.0	100.0	39953
Own Home (%)	97.0	100.0	16.9	100.0	100.0	100.0	100.0	41562
Number of Adults	2.0	2.0	1.6	-1.0	1.0	3.0	4.0	51955
Male (%)	87.5	100.0	33.1	0.0	100.0	100.0	100.0	45094
Income (000's)	88.1	62.5	64.1	25.0	45.0	112.5	250.0	45240
Number of Securities	2.7	1.0	3.9	0.0	1.0	3.0	6.0	51957
Portfolio Value (000's)	26.5	9.7	83.5	1.9	4.7	22.0	52.9	51957
Average Security Value (000's)	9.2	4.4	26.5	1.0	2.2	8.9	18.3	51957
Portfolio Return (%)	6.4	2.9	13.1	-0.5	1.0	7.1	17.0	49776
Sharpe Ratio (%)	12.4	14.6	46.0	-8.7	5.2	21.6	29.7	43109
Concentration	7.3	9.0	3.6	4.0	7.0	9.0	10.0	51957
Institutional Investors								
Portfolio Size ($m)	2230.8	290.8	13204.7	64.1	117.5	1004.8	3188.2	154952
HERF	0.1	0.0	0.1	0.0	0.0	0.1	0.2	154952
Institutions per county-quarter	5.0	1.3	19.4	1.0	1.0	3.4	8.6	299

Table 1
Summary Statistics (Continued...)

Panel B: ZIP code-level demographic characteristics (all U.S. ZIP codes)								
Variable	Mean	Median	Std Dev	10^{th} Pctl	25^{th} Pctl	75^{th} Pctl	90^{th} Pctl	N
Persons	8485.6	2822.0	12334.5	353.0	907.0	10756.0	26118.0	29305
Median Age	34.5	32.0	6.0	27.0	32.0	37.0	42.0	29305
Median Income (000's)	48.6	43.8	24.8	28.8	36.3	57.5	67.5	29305
Median Education (years)	12.6	12.5	1.0	12.5	12.5	12.5	13.5	29305
Minority (%)	11.6	2.9	19.1	0.0	0.6	13.5	36.8	29305
Married (%)	49.1	49.9	8.1	40.1	46.0	53.2	56.8	29305
Male-female Ratio	1.0	1.0	0.5	0.9	0.9	1.0	1.1	29305
Urban (%)	31.9	0.0	42.2	0.0	0.0	78.9	100.0	29305
Hispanic (%)	4.4	0.7	11.5	0.0	0.0	2.7	10.6	29305
Black (%)	7.1	0.4	15.8	0.0	0.0	4.7	23.8	29305

Panel C: ZIP code-level demographic characteristics (ZIP codes where retail investors reside)								
Variable	Mean	Median	Std Dev	10^{th} Pctl	25^{th} Pctl	75^{th} Pctl	90^{th} Pctl	N
Persons	17398.8	13741.0	15155.1	1667.0	5002.0	26045.0	38179.0	10484
Median Age	34.2	32.0	5.5	27.0	32.0	37.0	42.0	10484
Median Income (000's)	62.0	57.5	31.2	36.3	43.8	67.5	87.5	10484
Median Education (years)	13.0	12.5	0.9	12.5	12.5	13.5	13.5	10484
Minority (%)	13.1	6.1	17.6	0.7	2.0	16.5	35.9	10484
Married (%)	47.4	48.5	7.5	38.2	44.4	51.7	54.5	10484
Male-female Ratio	1.0	1.0	0.3	0.9	0.9	1.0	1.1	10484
Urban (%)	65.1	85.9	40.5	0.0	26.6	100.0	100.0	10484
Hispanic (%)	6.1	1.7	11.8	0.2	0.6	5.7	16.9	10484
Black (%)	7.4	1.6	14.4	0.0	0.3	6.7	21.2	10484
H/W	0.1	0.0	0.3	0.0	0.0	0.1	0.2	10484

Panel D: County-level demographic characteristics (all U.S. counties)								
Variable	Mean	Median	Std Dev	10^{th} Pctl	25^{th} Pctl	75^{th} Pctl	90^{th} Pctl	N
Persons	84855.3	23661.8	275367.2	5337.3	10799.5	59595.8	162428.8	3143
Median Age	36.5	37.0	4.1	31.0	34.5	38.9	40.9	3143
Education	0.1	0.1	0.1	0.1	0.1	0.2	0.2	3138
Married	0.6	0.6	0.1	0.5	0.6	0.6	0.7	3143
Male-female Ratio	1.0	1.0	0.1	0.9	0.9	1.0	1.0	3143
Urban (%)	15.9	1.9	38.3	0.1	0.3	12.6	46.5	3143
Hispanic (%)	5.7	1.7	11.4	0.6	0.9	4.4	13.6	3143
Black (%)	8.6	1.6	14.3	0.1	0.2	9.9	30.4	3143
H/W	0.1	0.0	0.8	0.0	0.0	0.1	0.2	3143

Panel E: County-level demographic characteristics (counties where institutional investors reside)								
Variable	Mean	Median	Std Dev	10^{th} Pctl	25^{th} Pctl	75^{th} Pctl	90^{th} Pctl	N
Persons	488477.1	296358.0	733425.4	64006.5	137602.8	603972.0	972924.8	299
Median Age	34.6	35.1	3.5	30.5	31.8	37.1	38.1	299
Education	0.2	0.2	0.1	0.1	0.2	0.3	0.3	299
Married	0.5	0.5	0.1	0.5	0.5	0.6	0.6	299
Male-female Ratio	1.0	1.0	0.0	0.9	0.9	1.0	1.0	299
Urban (%)	70.7	84.2	32.3	0.0	62.0	94.1	99.3	299
Hispanic (%)	7.8	4.2	9.5	1.2	2.0	9.2	21.0	299
Black (%)	11.4	7.1	12.9	0.7	2.0	16.5	28.0	299
H/W	0.1	0.1	0.3	0.0	0.0	0.1	0.4	299

Table 2

Hispanic Culture and Investment Preferences: OLS Regression Estimates

This table presents estimates from ordinary least squares regressions of the excess weight of investor i's portfolio on the set of stocks s at time t on a vector of local area and portfolio level covariates. The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include H/W, Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state and time fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. Columns (1) and (3) include ZIP code-level explanatory variables and household level regressors while columns (2) and (4) - (7) add MSA fixed effects to the model. We control for H/W, Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/W, Foreign, Density, $I_{Married}$, $I_{Own\ Home}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). The set of covariates is constant across columns (4) - (7). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors

	HQ < 60 Miles	Lottery Stocks	$Returns_{p>90}$	
	(1)	(2)	12 mo.	6 mo.
H/W	27.619	0.034	0.103	0.066
	(4.76)	(2.44)	(6.60)	(5.07)
Persons	-34.394	-0.166	-0.242	-0.171
	(-5.37)	(-10.28)	(-12.63)	(-12.24)
Urban	-49.591	0.055	0.089	0.056
	(-7.82)	(5.79)	(10.53)	(6.85)
Male/Female Ratio	-18.412	-0.014	-0.045	-0.041
	(-4.57)	(-0.79)	(-2.10)	(-2.45)
Median Age	-2.152	-0.022	-0.029	-0.038
	(-0.59)	(-1.81)	(-2.26)	(-3.37)
Port. Controls	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
State Ind.	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Adj. R sq.	0.176	0.068	0.065	0.048
N	150,213	151,027	151,027	151,027

Panel B: Retail Investors

	HQ < 60 Miles		Lottery Stocks		$Returns_{p>90}$	
	1	2	3	4	12 mo.	6 mo.
H/W	253.809	-19.166	15.294	9.113	2.606	2.413
	(8.21)	(-0.74)	(3.86)	(2.09)	(2.65)	(2.26)
Persons	16.150	52.864	-10.176	-4.526	0.529	1.358
	(0.71)	(2.86)	(-3.47)	(-1.48)	(0.69)	(1.61)
I_{Urban}	255.294	210.864	10.252	8.015	2.969	2.001
	(4.33)	(4.10)	(1.56)	(1.14)	(1.67)	(1.03)
I_{Male}	145.658	237.317	106.260	105.247	13.719	19.307
	(3.07)	(5.41)	(14.77)	(14.28)	(7.57)	(9.82)
Age	46.753	24.114	-3.615	-5.885	-10.624	-11.149
	(1.82)	(1.01)	(-1.17)	(-1.86)	(-13.70)	(-13.13)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
Adj. R sq.	0.007	0.101	0.010	0.015	0.011	0.008
N	1,508,936	1,460,572	1,470,297	1,409,559	1,409,559	1,409,559

Table 3

Hispanic Culture and Investment Preferences: IV Regression Estimates

This table presents estimates from instrumental variable regressions of the excess weight of investor i's portfolio on the set of stocks s at time t on a vector of local area and portfolio level covariates. The instrumental variable, MinDist, is the minimum distance from the Canadian border to investor i's locality (i.e., county or ZIP code). The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. Columns (1) and (3) include ZIP code-level explanatory variables and household level regressors while columns (2) and (4) - (7) add MSA fixed effects to the model. We control for Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/W, Foreign, Density, $I_{Married}$, $I_{Own\ Home}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). The set of covariates is constant across columns (4) - (7). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors

	HQ < 60 Miles	Lottery Stocks	$\text{Returns}_{p>90}$	
	(1)	(2)	12 mo.	6 mo.
H/W IV	53.986	0.314	0.235	0.383
	(2.24)	(3.53)	(2.29)	(4.03)
Persons	-47.113	-0.359	-0.364	-0.396
	(-2.62)	(-6.38)	(-5.68)	(-6.47)
Urban	-53.681	0.133	0.147	0.122
	(-8.56)	(13.98)	(15.78)	(14.01)
Male/Female Ratio	-23.138	0.099	0.006	0.027
	(-4.21)	(4.92)	(0.24)	(1.34)
Median Age	-1.434	0.071	0.004	0.014
	(-0.37)	(5.92)	(0.30)	(1.23)
Port. Controls	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
State Ind.	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	No	No
N	150,217	151,031	151,031	151,031

Panel B: Retail Investors

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{p>90}$	
	1	2	3	4	12 mo.	6 mo.
H/W IV	3737.564	3092.067	77.897	266.987	150.359	216.961
	(11.74)	(2.98)	(4.23)	(1.86)	(3.33)	(3.66)
Persons	-110.763	-1.415	-12.463	-8.908	-1.982	-2.288
	(-3.20)	(-0.04)	(-4.10)	(-2.08)	(-1.37)	(-1.19)
I_{Urban}	233.125	226.145	10.157	10.361	4.314	3.953
	(3.53)	(3.88)	(1.54)	(1.38)	(1.96)	(1.46)
I_{Male}	98.807	185.942	105.392	100.525	11.014	15.378
	(1.86)	(3.69)	(14.62)	(12.54)	(5.10)	(6.01)
Age	77.926	52.069	-3.033	-3.211	-9.092	-8.924
	(2.78)	(1.94)	(-0.98)	(-0.90)	(-9.22)	(-7.51)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
N	1,508,936	1,460,572	1,470,297	1,409,559	1,409,559	1,409,559

Table 4
Regression Estimates with Controls for Unobserved Geographic Heterogeneity

This table presents estimates from regressions of the excess weight ($EW_{i,s,t}$) of household i's portfolio on the set of stocks s at time t following the methodology of Rajan and Zingales (1998). The sample is from a discount brokerage account from January 1991 to November 1996. Panel A shows results for the H/W concentration variable while Panel B uses a ZIP code's minimum distance to the Canadian border, MinDist, as an IV for the Hispanic concentration of a ZIP code. The regressions include MSA-level fixed effects in an effort to further control for unobserved geographic heterogeneity. In addition, we include interaction terms of H/W and the minimum distance to the Canadian border with measures of household i's portfolio concentration. CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. In columns (1) and (2), the excess weight is measured with respect to the market weight of local stocks. A firm is defined to be "local" if it is headquartered within sixty miles of household i's ZIP code. Columns (3) and (4) present results where the dependent variable is the excess weight of household i's portfolio on lottery stocks. They are defined as firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. The remaining columns present results where the dependent variable is the excess weight of household i's portfolio on stocks in the top decile of returns over the last 12 and 6 month periods, respectively. The independent variable of interest in each model is the interaction of H/W with a measure of portfolio concentration (H/W×CONC or H/W×HERF). H/W is the ratio of Hispanic to White individuals in household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, $I_{Married}$ equals one if it is a married household, and $I_{Own\ Home}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). ZIP code-level controls are included but suppressed for brevity. Persons is the census count of individuals in household i's ZIP code. Persons is the census count of individuals in household i's ZIP code. B/W is the ratio of Black to White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in H/W. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

Panel A: OLS Regression

	HQ < 60		Lottery Stocks		$Returns_{p>90}$			
	1	2	3	4	12 mo.		6 mo.	
H/W x CONC	60.518		9.733		1.058		2.028	
	(4.04)		(4.51)		(2.22)		(3.89)	
H/W x HERF		659.226		176.286		14.426		24.625
		(2.90)		(6.37)		(2.22)		(3.47)
I_{Male}	237.649	248.443	105.489	104.820	13.775	13.032	19.367	18.408
	(5.42)	(5.49)	(14.32)	(14.05)	(7.60)	(7.06)	(9.85)	(9.21)
$I_{Married}$	-25.547	-30.937	-23.189	-25.826	-3.954	-3.427	-4.638	-3.953
	(-0.62)	(-0.72)	(-3.79)	(-4.13)	(-2.72)	(-2.31)	(-2.90)	(-2.43)
$I_{Own\ Home}$	215.705	197.530	-9.111	-5.595	-2.212	-2.408	-3.559	-3.838
	(4.11)	(3.60)	(-1.04)	(-0.63)	(-1.04)	(-1.11)	(-1.55)	(-1.64)
Age	28.383	37.596	-5.316	-5.007	-10.579	-10.872	-11.036	-11.415
	(1.19)	(1.53)	(-1.68)	(-1.55)	(-13.63)	(-13.77)	(-12.99)	(-13.26)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.101	0.104	0.015	0.016	0.011	0.012	0.008	0.008
N	1,460,572	1,354,339	1,409,559	1,329,347	1,409,559	1,329,347	1,409,559	1,329,347

Table 4
Regression Estimates with Controls for Unobserved Geographic Heterogeneity (Continued...)

Panel B: Instrumental Variable Regressions								
	HQ < 60		Lottery Stocks		$\text{Returns}_{\text{p} > 90}$			
	1	2	3	4	12 mo.		6 mo.	
H/W IV x CONC	383.935		44.694		2.019		5.286	
	(9.91)		(10.75)		(2.05)		(4.92)	
H/W IV x HERF		4978.536		1314.527		47.846		124.762
		(6.40)		(18.07)		(2.90)		(6.73)
I_{Male}	241.115	250.768	105.762	105.394	13.783	13.049	19.392	18.458
	(5.48)	(5.51)	(14.34)	(13.93)	(7.60)	(7.07)	(9.87)	(9.22)
I_{Married}	-21.873	-21.368	-22.781	-23.427	-3.943	-3.356	-4.600	-3.742
	(-0.53)	(-0.50)	(-3.72)	(-3.71)	(-2.71)	(-2.27)	(-2.88)	(-2.30)
$I_{\text{Own Home}}$	200.328	175.927	-10.842	-11.595	-2.259	-2.584	-3.720	-4.365
	(3.81)	(3.19)	(-1.23)	(-1.28)	(-1.06)	(-1.19)	(-1.62)	(-1.86)
Age	50.279	56.300	-2.934	0.005	-10.513	-10.724	-10.813	-10.974
	(2.11)	(2.31)	(-0.92)	(0.00)	(-13.53)	(-13.58)	(-12.72)	(-12.72)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,460,572	1,354,339	1,409,559	1,329,347	1,409,559	1,329,347	1,409,559	1,329,347

Table 5
Investor Herding Estimates

This table tests whether the trades of investors are correlated. We calculate the herding measure for firms local to investors, lottery stocks, and high-momentum stocks in each H/W quintile. A firm is local to an investor if it is headquartered within 60 miles of investors i's location. Lottery stocks include firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. High-momentum stocks include those in the top decile of returns over the most recent 12 and 6 month periods, respectively. For Panel A, we use ANcerno Ltd. data which contains institutional trading data from January 1999 to December 2010. The sample for Panel B consists of monthly trades made by retail investors from January 1991 to November 1996 at a large discount brokerage house. The herding measure for stock i in month t is the following: $HM_{i,t} = |p_{i,t} - E[p_{i,t}]| - E|p_{i,t} - E[p_{i,t}]|$ (Lakonishok, Shleifer, and Vishny (1992); Barber, Odean, and Zhu (2009b)). $p_{i,t}$ is the proportion of purchases to the total number of trades in stock i during month t. $E[p_{i,t}]$ is the proportion of purchases to the total number of trades in month t. $|p_{i,t} - E[p_{i,t}]|$ is the proportion of purchases to the total number of trades in stock i during month t minus the proportion of all purchases during month t. The latter term of the measure, $E|p_{i,t} - E[p_{i,t}]|$, is used to adjust for the fact that more variation in the proportion of buys is expected in stocks that have a few trades. We restrict the analysis to stocks with at least ten trades in month t. In each month, we average the herding measures across stocks. Statistical tests are based on the time-series of the mean herding measure across stocks. *p-values* can be found in parentheses. We test the differences between the high- and low-Hispanic herding measures in the column labeled Difference and report the p-values.

Panel A: Institutional Investors

	High HW	*p-value*	Low HW	*p-value*	Difference	*p-value*
Firms Local to Investors	0.246	(0.00)	0.193	(0.00)	0.053	(0.00)
Lottery Stocks	0.250	(0.00)	0.234	(0.00)	0.016	(0.00)
Momentum - 12 mo.	0.240	(0.00)	0.222	(0.00)	0.018	(0.00)
Momentum - 6 mo.	0.227	(0.00)	0.218	(0.00)	0.009	(0.12)

Panel B: Retail Investors

	High HW	*p-value*	Low HW	*p-value*	Difference	*p-value*
Local Firms	0.058	(0.00)	-0.022	(0.88)	0.081	(0.00)
Lottery Stocks	0.017	(0.04)	0.016	(0.07)	0.001	(0.92)
Momentum - 12 mo.	0.039	(0.00)	0.017	(0.03)	0.022	(0.05)
Momentum - 6 mo.	0.041	(0.00)	0.023	(0.01)	0.018	(0.16)

Table 6
Hispanic Concentration and Return Comovements

This table shows the results for the comovement of returns tests. For each H/W quintile, we estimate a time-series regression of monthly stock returns on the returns of the habitat portfolio along with several controls. Panel A includes the market portfolio, the Fama and French (1993) three-factors, the Carhart (1997) momentum and Pastor and Stambaugh (2003) liquidity factor. Panel B includes the market portfolio, the Fama and French (1993) three-factors, short-term and long-term reversals. Cross-sectional averages of the estimate coefficients (habitat betas) from the time-series regressions and their t-statistics are presented in the table. The habitat portfolio is constructed as a value-weighted portfolio of all local, lottery, and high-momentum firms found in the same H/W quintile as firm i, excluding firm i. The market index is the value-weighted return of all stocks in the market minus firm i and the firms included in the habitat portfolio. Local firms are defined as the companies headquartered within 60 miles of an investor's ZIP code. Lottery stocks are firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. High-momentum stocks are the firms in the top decile of returns over the most recent 12 and 6 month periods. Each panel also includes a within quintile index that measures how well each stock meets these characteristics: being local, have a high degree of lotteriness, and experience high-momentum returns. We then divide stocks into quintiles using this index. The column Diff. High H/W and Low H/W shows the difference between high and low-Hispanic betas. The row *Diff. High and Low Index* tests whether the difference between the betas of "High Index" and "Low Index" quintiles are statistically significant. The p-values of the differences are reported in the row below in between parentheses. Standard errors are adjusted using the Newey and West (1987) correction method.

Panel A: Controlling for the Fama-French 3 Factors, Momentum, and Liquidity

	High H/W	4	3	2	Low H/W	Diff. High and Low H/W
Firms Local to Investors	0.576	0.234	0.045	0.453	0.120	0.456
	(10.14)	(4.83)	(0.79)	(12.09)	(2.08)	(0.00)
Lottery Stocks	0.838	0.655	0.630	0.579	0.321	0.517
	(13.45)	(10.86)	(9.28)	(10.68)	(5.27)	(0.00)
Momentum - 12 mo.	0.467	0.283	0.162	0.185	0.051	0.416
	(11.83)	(4.65)	(3.27)	(5.31)	(1.43)	(0.00)
Momentum - 6 mo.	0.437	0.159	0.149	0.160	0.078	0.359
	(14.04)	(3.84)	(3.74)	(4.51)	(1.90)	(0.00)
High Index	0.501				0.135	0.366
	(9.78)				(3.43)	(0.00)
4	0.403				0.077	0.326
	(7.80)				(1.61)	(0.00)
3	0.347				0.149	0.198
	(4.92)				(2.88)	(0.04)
2	0.341				0.165	0.176
	(5.17)				(3.63)	(0.03)
Low Index	0.224				0.080	0.144
	(4.00)				(2.10)	(0.03)
Diff. High and Low Index	0.277				0.055	
	(0.00)				(0.31)	

Table 6
Hispanic Concentration and Return Comovements (Continued...)

Panel B: Controlling for the Fama-French 3 Factors, Short-term and Long-term Reversal						
	High H/W	4	3	2	Low H/W	Diff. High and Low H/W
Firms Local to Investors	0.576	0.229	0.030	0.449	0.125	0.451
	(11.22)	(4.65)	(0.56)	(12.51)	(2.28)	(0.00)
Lottery Stocks	0.823	0.640	0.607	0.565	0.312	0.511
	(13.30)	(11.79)	(9.33)	(10.94)	(5.21)	(0.00)
Momentum - 12 mo.	0.463	0.255	0.157	0.174	0.044	0.419
	(13.60)	(4.83)	(3.31)	(5.00)	(1.23)	(0.00)
Momentum - 6 mo.	0.439	0.158	0.148	0.157	0.076	0.363
	(14.29)	(3.83)	(3.92)	(4.52)	(1.83)	(0.00)
High Index	0.494				0.131	0.363
	(10.83)				(3.35)	(0.00)
4	0.400				0.065	0.335
	(7.84)				(1.44)	(0.01)
3	0.344				0.158	0.186
	(5.25)				(3.19)	(0.04)
2	0.337				0.163	0.174
	(5.17)				(3.68)	(0.03)
Low Index	0.223				0.084	0.139
	(3.96)				(2.15)	(0.04)
Diff. High and Low Index	0.271				0.047	
	(0.00)				(0.40)	

Table 7
Hispanic Concentration and Local Stock Returns

This table reports performance estimates of trading strategies for local stocks. We perform a bivariate dependent sort where we first sort stocks into quintiles based on their H/W MSA value. We then sort stocks into quintiles based on their previous quarter change in local institutional ownership (IO). Panel A presents mean monthly value-weighted excess returns (i.e., in excess of the risk-free rate). The *H-L* row and Difference column contain point estimates and t-statistics from a test of equality of returns between low and high H/W location companies and low and high IO firms, respectively. Panel B presents the performance estimates of the trading strategy. The zero-cost strategy consists of going long on the top quintile based on previous quarter change in local IO and short on the bottom quintile based on previous quarter change in local IO, for each H/W percentile. We regress the portfolio returns on a set of factors, which include the market excess return (RMRF), size (SMB), value (HML), short-term reversal factor (STR) and long-term reversal factor (LTR). t-statistics are presented in parentheses and are adjusted for autocorrelation and heteroskedasticity following Newey and West (1987). The Alpha Difference row is the difference in the alphas of the trading strategies between the high and low H/W areas. The p-value testing equality of alphas is presented in parentheses.

Panel A: VW Excess Return				*Panel B: Trading Strategies Performance*			
	Low ΔIO	High ΔIO	Difference		Low ΔIO (S)	High ΔIO (L)	Long-Short
Low H/W	0.772	0.696	-0.076	Low H/W	0.025	0.014	-0.011
	(2.69)	(2.65)	(-0.35)		(0.16)	(0.10)	(-0.05)
2	0.687	0.928	0.241	2	-0.006	0.273	0.279
	(2.55)	(3.09)	(0.87)		(-0.03)	(1.55)	(1.02)
3	0.364	0.715	0.351	3	-0.401	0.116	0.516
	(1.17)	(2.69)	(1.40)		(-2.03)	(0.86)	(2.19)
4	0.684	0.817	0.133	4	-0.137	0.145	0.282
	(1.81)	(2.16)	(0.40)		(-0.62)	(0.58)	(0.88)
High H/W	0.261	1.060	0.799	High H/W	-0.519	0.431	0.950
	(0.85)	(2.81)	(2.56)		(-2.55)	(1.73)	(2.86)
H-L			0.875	Alpha Difference			0.961
			(2.54)	(5-1)			(0.00)

Table 8
Hispanic Concentration and Lottery-Stock Premium

This table reports performance estimates for lottery and non-lottery stocks in low and in high H/W MSAs. Panel A presents mean monthly value-weighted excess returns (i.e., in excess of the risk-free rate). The *H-L* row and Difference column contain point estimates and t-statistics from a test of equality of returns between low and high H/W location companies. Panel B reports the risk-adjusted performance estimates for lottery and non-lottery stocks in low and in high H/W MSAs. The set of factors for the portfolios includes the market excess return (RMRF), size (SMB), value (HML), short-term reversal factor (STR) and long-term reversal factor (LTR). The *L-S* row captures the difference in the returns of the Long and Short portfolios. t-statistics are presented in parentheses and are adjusted for autocorrelation and heteroskedasticity following Newey and West (1987). The Alpha Difference column is the difference in the alphas of the trading strategies between lottery and non-lottery stocks for each H/W percentile. The p-value testing equality of alphas is presented in parentheses.

Panel A: VW Excess Return				*Panel B: Portfolio Performance*			
	Lottery	Non-Lottery	Difference		Lottery	Non-Lottery	Alpha Difference
Low H/W	-1.876	0.471	-2.347	High H/W (S)	-2.981	-0.100	-2.881
	(-4.22)	(2.52)	(-6.49)		(-9.65)	(-1.16)	(0.00)
2	-1.957	0.641		2	-3.001	0.048	
	(-4.69)	(3.41)			(-12.79)	(0.68)	
3	-2.165	0.607		3	-3.154	0.106	
	(-4.95)	(2.89)			(-13.14)	(1.74)	
4	-2.287	0.497		4	-3.275	-0.041	
	(-4.73)	(2.67)			(-11.45)	(-0.80)	
High H/W	-2.425	0.536	-2.961	High H/W (L)	-3.511	0.046	-3.557
	(-5.24)	(2.66)	(-11.32)		(-13.98)	(1.30)	(0.00)
H-L	-0.549	0.065	-0.614	L-S	-0.531	0.146	-0.677
	(-1.94)	(0.59)	(-2.10)		(-2.17)	(1.36)	(0.02)

Table 9
Hispanic Concentration and Momentum Returns

This examines whether momentum returns are higher for stocks headquartered in high-Hispanic areas. Panel A reports mean monthly returns for a portfolio of winners, a portfolio of losers and a winners-minus-losers momentum portfolio, by Hispanic concentration (H/W) in the MSA in which the company is headquartered. We sort all stocks into quintiles based on the ratio of Hispanic population to white population in the MSA in which the company is headquartered, according to the decennial census. We then sort the MSA-level portfolios into winners and losers. "Winners" are those companies with stock returns in the highest decile in the (t-7, t-1) period, with a one-month delay in portfolio formation to avoid the short-term reversal phenomenon. "Losers" are those companies with stock returns in the lowest decile in the (t-7, t-1) period. *t*-statistics are presented in parentheses and are corrected for heteroskedasticity and serial correlation using the method of Newey and West (1987). Point estimates and *t*-statistics from a test of equality of returns between low and high H/W location companies are presented in the last two rows. Panel B presents the risk-adjusted performance estimates for the winner-minus-loser momentum strategy in low and in high H/W MSAs. Component returns are those of equally weighted portfolios of companies in high and low H/W MSAs. The set of factors includes the market excess return (RMRF), size (SMB), value (HML), short-term reversal factor (STR) and long-term reversal factor (LTR). *t*-statistics are presented in parentheses and are adjusted for autocorrelation and heteroskedasticity following Newey and West (1987). Alpha Difference is the difference in the alpha of the momentum strategy between the high and low H/W MSAs. The *p*-value testing equality of alphas is presented in parentheses. The estimation period is January 1970 to December 2018.

Panel A: Momentum Returns

	Winners (W)		Losers (L)		Momentum Portfolio (W-L)
	Raw Return	Mean H/W	Raw Return	Mean H/W	Raw Return
Low H/W	1.042	0.01	0.83	0.01	0.212
	(4.01)		(2.41)		(0.86)
2	1.254	0.03	0.58	0.03	0.674
	(4.63)		(1.80)		(3.25)
3	1.159	0.07	0.31	0.07	0.849
	(3.83)		(0.94)		(3.67)
4	0.997	0.14	0.452	0.14	0.544
	(3.45)		(1.32)		(2.40)
High H/W	1.115	0.29	0.328	0.29	0.787
	(3.49)		(0.92)		(3.27)
H-L	0.073		-0.502	MOM_5-MOM_1	0.575
	(0.18)		(-1.04)		(1.74)

Table 9
Hispanic Concentration and Momentum Returns (Continued...)

Panel B: Risk-adjusted Momentum Returns						
	Low H/W	High H/W	Low H/W	High H/W	Low H/W	High H/W
Alpha	0.253	0.851	0.362	1.013	0.433	1.134
	(1.08)	(3.83)	(1.52)	(4.42)	(1.77)	(4.68)
RMRF	-0.077	-0.121	-0.131	-0.187	-0.081	-0.107
	(-1.08)	(-1.36)	(-1.88)	(-2.33)	(-0.94)	(-1.07)
SMB			0.051	0.006	0.062	-0.04
			(0.39)	(0.03)	(0.51)	(-0.24)
HML			-0.255	-0.374	-0.248	-0.458
			(-2.21)	(-1.89)	(-1.69)	(-2.05)
STR					-0.242	-0.413
					(-1.81)	(-2.28)
LTR					0.046	0.282
					(0.33)	(1.46)
Alpha Difference	0.598		0.651		0.701	
(5-1)	(0.004)		(0.003)		(0.001)	
Adj. R sq.	0.002	0.007	0.017	0.036	0.032	0.089
N	588	588	588	588	588	588

Figure A1
Hispanic Concentration in the U.S., by County

This figure plots the Hispanic concentration of a county as measured by H/W, the ratio of the Hispanic population to the White population in a county. Concentration is measured at the beginning of each decade based on the most recent census and is held constant for that decade. The decennial censuses used in this study start in 1970 and end in 2010. All of the counties in the U.S. are included.

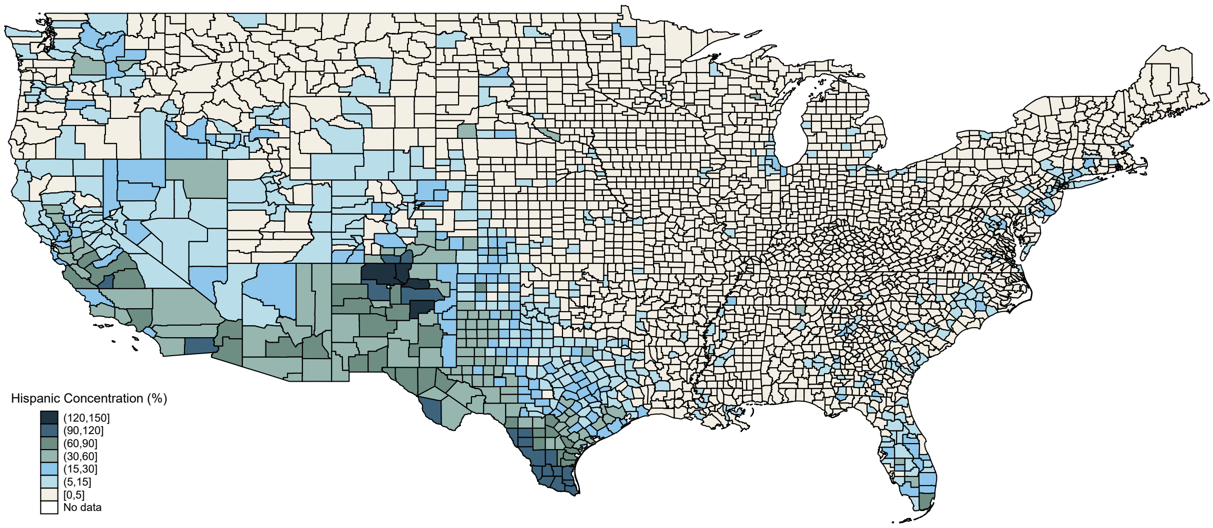

Figure A2
Counties Along the Canadian Border

This figure plots the counties along the Canadian border used to calculate the minimum distance to each ZIP code in the U.S.

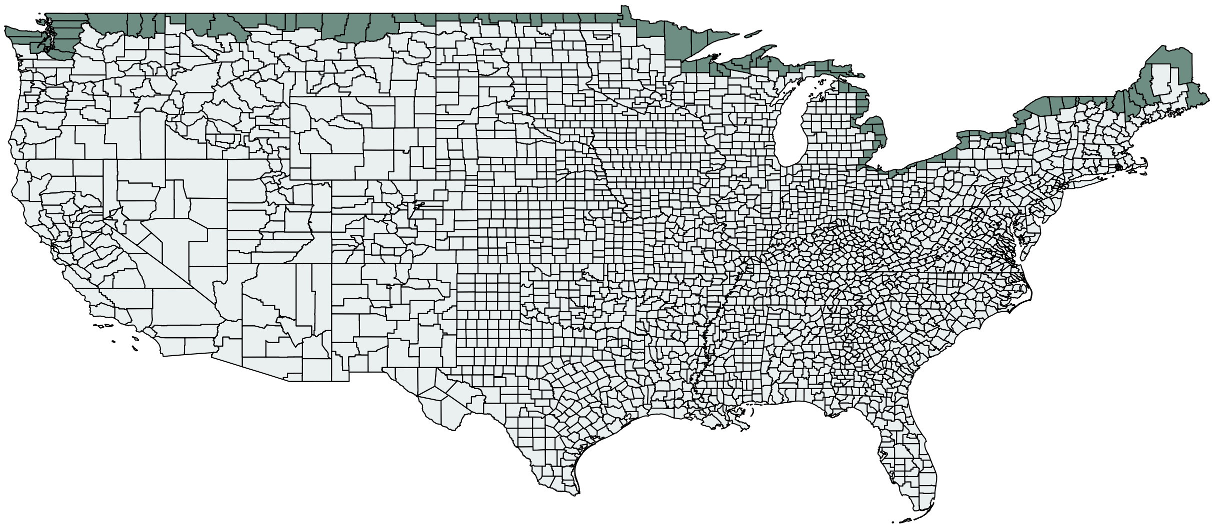

Table A1
Variable Definitions

Institutional Investor Variables	
Variable	Definitions
H/W	The concentration of Hispanics in an area. The ratio of Hispanic to White individuals in the investor's county.
Persons	The total county-level population.
Urban	The proportion of the county population that lives in urban areas.
Male-female Ratio	The ratio of male to female residents in a county.
Median Age	The median age of county residents.
Married	The proportion of county households with a married household.
Education	The proportion of county-level population over the age of 25 with a bachelor's degree or higher.
Portfolio Value	The market value of the total institutional portfolio.
HERF	The Herfindahl Index of the institution's portfolio.
Hispanic	The percentage of people identifying as having a Hispanic ancestry in each county.
Black	The percentage of people identifying as having a Black ancestry in each county.
H/P	The ratio of Hispanic to total individuals in the investor's county.
B/P	The ratio of Black to total individuals in the investor's county.
MinDist	The minimum distance from the Canadian border to investor's county.

Retail Investor Variables	
Variable	Definitions
H/W	The concentration of Hispanics in an area. The ratio of Hispanic to White individuals in the investor's ZIP code.
Age	The age of the head of the household.
I_{Married}	Indicator variable equal to one if the head of the household is married, and zero otherwise.
$I_{\text{Own Home}}$	Dummy variable equal to one if the household owns it home, and zero otherwise.
I_{Male}	Indicator variable equal to one if the head of the household is male, and zero otherwise.
Level of Education	The highest educational level of the head of the household.
Income	The household's income category. There are nine categories with the following midpoints (in thousands): 7.5, 17.5, 25, 35, 45, 62.5, 87.5, 112.5, and 250.
Knowledge	A household's self-reported financial knowledge, measured on a zero-to-four scale.
Experience	A household's self-reported financial experience, measured on a zero-to-four scale.
Number of Securities	The number of securities owned by the household in their brokerage account.
Portfolio Value	The total value of the portfolio (in thousands).
Average Security Value	The average dollar value (in thousands) of all securities held in the brokerage account.
Portfolio Return	The portfolio average annual return.
Sharpe Ratio	The portfolio average Sharpe Ratio.
Concentration	Measure of the household's portfolio concentration. It is defined as ten minus the number of stocks in the portfolio.
HERF	The Herfindahl Index for the portfolios of all long-only investors.
Persons	The total ZIP code-level population.
Median Age	The median age of the residents in a ZIP code.
Median Income	The median household income (in thousands) in a ZIP code.
Median Education	The median level of education (in years) in each ZIP code.

Table A1
Variable Definitions (Continued...)

Retail Investor Variables	
Variable	Definitions
Minority	The percentage of the population classified as belonging to a minority (non-white) in each ZIP code.
Married	The proportion of of households in a ZIP code that are married.
Male-female Ratio	The ratio of men to women in a ZIP code.
Urban	The proportion of the ZIP code population that lives in urban areas.
Hispanic	The percentage of people identifying as having a Hispanic ancestry in each ZIP code.
Black	The percentage of people identifying as having a Black ancestry in each ZIP code.
H/P	The ratio of Hispanic to total individuals in the investor's ZIP code.
B/P	The ratio of Black to total individuals in the investor's ZIP code.
MinDist	The minimum distance from the Canadian border to investor's ZIP code.

Table A2
Counties Along the Canadian Border

This table presents the counties along the Canadian border used to calculate the minimum distance to each ZIP code in the U.S.

State	*County*	*FIPS*	*State*	*County*	*FIPS*	*State*	*County*	*FIPS*
Washington	Clallam	53009	North Dakota	Pembina	38067	Michigan	Wayne	26163
Washington	Jefferson	53031	Minnesota	Kittson	27069	Michigan	Monroe	26115
Washington	Mason	53045	Minnesota	Roseau	27135	Ohio	Lucas	39095
Washington	Pierce	53053	Minnesota	Lake of the Woods	27077	Ohio	Ottawa	39123
Washington	Kitsap	53035	Minnesota	Koochiching	27071	Ohio	Eerie	39043
Washington	King	53033	Minnesota	St. Louis	27137	Ohio	Lorain	39093
Washington	Snohomish	53061	Minnesota	Lake	27075	Ohio	Cuyahoga	39035
Washington	Island	53029	Minnesota	Cook	27031	Ohio	Lake	39085
Washington	San Juan	53055	Wisconsin	Douglas	550331	Ohio	Ashtabula	39007
Washington	Skagit	53057	Wisconsin	Bayfield	55007	Pennsylvania	Eerie	42049
Washington	Whatcom	53073	Wisconsin	Ashland	55003	New York	Chautauqua	36013
Washington	Okanogan	53047	Wisconsin	Iron	55051	New York	Eerie	36029
Washington	Ferry	53019	Michigan	Gogebic	26053	New York	Niagara	36063
Washington	Stevens	53065	Michigan	Ontonagon	26131	New York	Orleans	36073
Washington	Pend Oreille	53051	Michigan	Houghton	26061	New York	Monroe	36055
Idaho	Boundary	16021	Michigan	Keweenaw	26083	New York	Wayne	36117
Montana	Lincoln	30053	Michigan	Baraga	26013	New York	Cayuga	36011
Montana	Flathead	30029	Michigan	Marquette	26103	New York	Oswego	36075
Montana	Glacier	30035	Michigan	Alger	26003	New York	Jefferson	36045
Montana	Toole	30101	Michigan	Luce	26095	New York	St. Lawrence	36089
Montana	Liberty	30051	Michigan	Chippewa	26033	New York	Franklin	36033
Montana	Hill	30041	Michigan	Mackinac	26097	New York	Clinton	36019
Montana	Blaine	30005	Michigan	Cheboygan	26031	Vermont	Grand Isle	50013
Montana	Phillips	30071	Michigan	Presque	26141	Vermont	Franklin	50009
Montana	Valley	30105	Michigan	Alpena	26007	Vermont	Orleans	50019
Montana	Daniels	30019	Michigan	Alcona	26001	Vermont	Essex	50011
Montana	Sheridan	30091	Michigan	Iosco	26069	New Hampshire	Coos	33007
North Dakota	Divide	38023	Michigan	Arenac	26011	Maine	Oxford	23017
North Dakota	Burke	38013	Michigan	Bay	26017	Maine	Franklin	23007
North Dakota	Renville	38075	Michigan	Tuscola	26157	Maine	Somerset	23025
North Dakota	Bottineau	38009	Michigan	Huron	26063	Maine	Aroostook	23003
North Dakota	Rolette	38079	Michigan	Sanilac	26151	Maine	Washington	23029
North Dakota	Towner	38095	Michigan	St. Clair	26147			
North Dakota	Cavalier	38019	Michigan	Macomb	26099			

Table A3

Univariate Regressions: Hispanic Culture and their Investment Preferences

This table presents univariate regressions of the excess weight of investor i's portfolio on the set of stocks s at time t. The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. In the first column, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second column, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last two columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors					*Panel B: Retail Investors*				
	HQ < 60 Miles	Lottery Stocks	Returns$_{\text{p}}$ > 90			HQ < 60 Miles	Lottery Stocks	Returns$_{\text{p}}$ ¿ 90	
	(1)	(2)	12 mo.	6 mo.		1	2	12 mo.	6 mo.
H/W	5.351	0.189	0.179	0.147	H/W	-17.675	22.509	1.540	4.002
	(3.02)	(11.33)	(11.86)	(10.99)		(-0.82)	(7.69)	(2.25)	(5.25)
Controls	No	No	No	No	Port. Chars	No	No	No	No
State Ind.	No	No	No	No	Income	No	No	No	No
Year-Quarter FE	No	No	No	No	MSA Ind.	No	No	Yes	Yes
Adj. R sq.	0.000	0.004	0.003	0.003	Adj. R sq.	0.000	0.000	0.000	0.000
N	150,213	151,027	151,027	151,027	N	1,508,936	1,470,297	1,409,559	1,409,559

Table A4
Univariate Regressions: Controlling for Unobserved Geographic Heterogeneity

This table presents univariate regressions of the excess weight ($EW_{i,s,t}$) of household i's portfolio on the set of stocks s at time t following the methodology of Rajan and Zingales (1998). We include interaction terms of H/W with measures of household i's portfolio concentration. CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. In columns (1) and (2), the excess weight is measured with respect to the market weight of local stocks. A firm is defined to be "local" if it is headquartered within sixty miles of household i's ZIP code. Columns (3) and (4) present results where the dependent variable is the excess weight of household i's portfolio on lottery stocks. They are defined as firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. The remaining columns present results where the dependent variable is the excess weight of household i's portfolio on stocks in the top decile of returns over the last 12 and 6 month periods, respectively. The independent variable of interest in each model is the interaction of H/W with a measure of portfolio concentration (H/W×CONC or H/W×HERF). H/W is the ratio of Hispanic to White individuals in household i's ZIP code. Note that interaction terms are not standardized to aid in comparisons across differences in H/W. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

	HQ < 60 Miles		Lottery Stocks		$Returns_{p > 90}$			
	1	2	3	4	12 mo.		6 mo.	
H/W x CONC	43.538		22.416		0.362		2.483	
	(3.18)		(11.23)		(0.83)		(5.10)	
H/W x HERF		389.264		261.408		7.007		27.616
		(2.23)		(10.88)		(1.31)		(4.25)
MSA Ind.	No	No	No	No	No	No	No	No
Income Controls	No	No	No	No	No	No	No	No
ZIP Code Controls	No	No	No	No	No	No	No	No
Portfolio Chars.	No	No	No	No	No	No	No	No
Adj. R sq.	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000
N	1,509,087	1,399,484	1,470,511	1,386,925	1,470,511	1,386,925	1,470,511	1,386,925

Table A5
First Stage Regressions: Hispanic Culture and their Investment Preferences

This table presents univariate and first stage regressions of an area's minimum distance to the Canadian border on the local Hispanic concentration measure. The instrumental variable, MinDist, is the minimum distance from the Canadian border to investor i's locality (i.e., county or ZIP code). The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. We control for Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/W, Foreign, Density, I_{Married}, $I_{\text{Own Home}}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{p>90}$			
	(1)	(2)	(3)	(4)	12 mo.		6 mo.	
MinDist	0.213	0.080	0.000	0.001	0.000	0.001	0.000	0.001
	(5.80)	(1.94)	(1.03)	(4.95)	(2.88)	(3.29)	(2.94)	(5.34)
Persons		-21.334		-0.174		-0.203		-0.164
		(-4.18)		(-12.85)		(-12.88)		(-13.01)
Urban		-49.630		0.056		0.089		0.056
		(-7.83)		(5.91)		(10.63)		(7.00)
Male/Female Ratio		-17.756		-0.014		-0.043		-0.041
		(-4.30)		(-0.81)		(-2.03)		(-2.44)
Median Age		-2.691		-0.014		-0.028		-0.032
		(-0.80)		(-1.20)		(-2.17)		(-2.89)
Port. Controls	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	No	Yes	No	Yes	No	Yes	No	Yes
State Ind.	No	Yes	No	Yes	No	Yes	No	Yes
Year-Quarter FE	No	No	No	No	No	No	No	No
Adj. R sq.	0.010	0.176	0.000	0.068	0.000	0.065	0.000	0.048
F-Stat.		354.19		320.96		320.96		320.96
N	150,213	150,213	151,027	151,027	151,027	151,027	151,027	151,027

Table A5
First Stage Regressions: Hispanic Culture and their Investment Preferences (Continued...)

	Panel B: Retail Investors											
	HQ < 60 Miles				Lottery Stocks				$\text{Returns}_{\text{p} > 90}$			
	1	2	3	4	5	6	7	8	12 mo.	12 mo.	6 mo.	6 mo.
Min. Dist.	0.001	0.000	0.002	0.001	0.001	0.000	0.002	0.001	0.002	0.001	0.002	0.001
	(31.51)	(24.90)	(4.19)	(4.31)	(32.29)	(25.42)	(4.91)	(4.89)	(4.91)	(4.89)	(4.91)	(4.89)
Persons		0.121		0.122		0.118		0.119		0.119		0.119
		(9.09)		(8.99)		(9.36)		(9.24)		(9.24)		(9.24)
I_{Urban}		0.528		0.559		0.531		0.561		0.561		0.561
		(41.41)		(33.90)		(41.42)		(33.77)		(33.77)		(33.77)
I_{Male}		0.017		0.017		0.017		0.018		0.018		0.018
		(2.61)		(2.58)		(2.63)		(2.88)		(2.88)		(2.88)
Age		-0.013		-0.008		-0.013		-0.010		-0.010		-0.010
		(-4.33)		(-2.95)		(-4.53)		(-3.42)		(-3.42)		(-3.42)
Port. Chars	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Income	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
MSA Ind.	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.072	0.476	0.196	0.532	0.074	0.472	0.205	0.536	0.205	0.536	0.205	0.536
F-Stat.		620.01		18.58		645.94		23.88		23.88		23.88
N	1,508,936	1,508,936	1,460,572	1,460,572	1,470,297	1,470,297	1,409,559	1,409,559	1,409,559	1,409,559	1,409,559	1,409,559

Table A6
First Stage Regressions: Controlling for Unobserved Geographic Heterogeneity

This table presents univariate and first stage regressions of a ZIP code's minimum distance to the Canadian border interacted with a measure of portfolio concentration (H/W×CONC or H/W×HERF). CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. The independent variable of interest in each model is the interaction of MinDist with a measure of portfolio concentration (MinDist×CONC or MinDist×HERF). MinDist is the minimum distance from the Canadian border to household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, I_{Married} equals one if it is a married household, and $I_{\text{Own Home}}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). ZIP code-level controls are included but suppressed for brevity. Persons is the census count of individuals in household i's ZIP code. Persons is the census count of individuals in household i's ZIP code. B/W is the ratio of Black to White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in MinDist. All specifications include MSA-level fixed effects. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

2-SLS First Stage Regressions								
	HQ < 60 Miles				Lottery Stocks			
	1	2	3	4	5	6	7	8
MinDist x CONC	0.017	0.017			0.017	0.017		
	(41.58)	(38.67)			(41.74)	(38.94)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(50.79)	(52.13)			(50.72)	(52.26)
I_{Male}		0.004		0.001		0.009		0.001
		(0.41)		(0.99)		(0.86)		(1.22)
I_{Married}		-0.017		-0.002		-0.019		-0.002
		(-1.98)		(-2.96)		(-2.15)		(-3.05)
$I_{\text{Own Home}}$		0.057		0.006		0.060		0.006
		(4.14)		(4.16)		(4.31)		(4.46)
Age		-0.022		-0.001		-0.023		-0.001
		(-4.67)		(-3.08)		(-4.77)		(-3.24)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.241	0.420	0.220	0.446	0.243	0.419	0.223	0.446
F-Stat.		1495.62		2717.46		1516.62		2731.29
N	1,460,572	1,460,572	1,354,339	1,354,339	1,409,559	1,409,559	1,329,347	1,329,347

Table A6

First Stage Regressions: Controlling for Unobserved Geographic Heterogeneity (Continued...)

	2-SLS First Stage Regressions							
	$\text{Returns}_{\text{p} > 90}$							
	12 mo.	12 mo.	12 mo.	12 mo.	6 mo.	6 mo.	6 mo.	6 mo.
MinDist x CONC	0.017	0.017			0.017	0.017		
	(41.74)	(38.94)			(41.74)	(38.94)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(50.72)	(52.26)			(50.72)	(52.26)
I_{Male}		0.009		0.001		0.009		0.001
		(0.86)		(1.22)		(0.86)		(1.22)
I_{Married}		-0.019		-0.002		-0.019		-0.002
		(-2.15)		(-3.05)		(-2.15)		(-3.05)
$I_{\text{Own Home}}$		0.060		0.006		0.060		0.006
		(4.31)		(4.46)		(4.31)		(4.46)
Age		-0.023		-0.001		-0.023		-0.001
		(-4.77)		(-3.24)		(-4.77)		(-3.24)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.243	0.419	0.223	0.446	0.243	0.419	0.223	0.446
F-Stat.		1516.62		2731.29		1516.62		2731.29
N	1,409,559	1,409,559	1,329,347	1,329,347	1,409,559	1,409,559	1,329,347	1,329,347

Internet Appendix to

Hispanic Culture, Stock Preferences, and Asset Prices

August 24, 2020

IA1 Hispanic Culture and Real Estate Market Returns

We study the U.S. residential real estate market to provide additional evidence that the asset pricing effects we have identified are driven by local investors living in high-Hispanic areas. The choice of the real estate market is justified by the fact that it provides a tighter link between the buyers of an asset and the asset itself. For instance, when someone buys a house, it is likely that they intend to live in that residence. Thus, due to the immovable nature of residential real estate, the local ownership level is high. The return chasing and preference for commonly used products in Hispanic communities implies that as more and more families buy real estate in areas with a high concentration of Hispanic residents, more people will find owning an asset in this area desirable, potentially resulting in a run-up in prices. We expect areas with a large Hispanic population to have more pronounced price run-ups and subsequent downturns in housing prices, and thus, higher realized volatility.

To test this hypothesis, we use CBSA-level housing price indices from the Federal Housing Finance Agency. Similar to the previous analysis, we sort core based statistical areas (CBSAs) into quintiles based on the concentration of Hispanics in the local population. We then calculate the growth in house prices by equally weighting each CBSA in a quintile.

Figure IA1 graphs the cumulative returns of housing prices for the 100 CBSAs in the U.S.[IA1] It shows that CBSAs with the highest concentration of Hispanic residents exhibit the largest swings in real estate prices. Figure IA2 graphs the raw returns for the different quintiles of Hispanic concentration and suggests that areas with a higher Hispanic concentration exhibit more variance in annual returns. A statistical test of equality of variance in house price returns between the lowest and highest Hispanic concentration quintiles is rejected at the 1% level, confirming that the variance of the 5$^{\text{th}}$ quintile is statistically larger.

The findings from the real estate tests confirm our conjecture: localities with a high-Hispanic population exhibit higher price run-ups and subsequent downturns in asset prices, as well as realized volatility. Furthermore, the results show that the Hispanic culture can affect two asset markets in the U.S., the equity and real estate markets.

[IA1]The data set is from the Federal Housing Finance Agency's House Price Index, which is based on transaction data for single family homes.

IA2 Alternative IV: Distance from the Mexican Border

We use a ZIP code's minimum distance to the Canadian border as our primary instrument for the Hispanic concentration of an area. However, we also test whether the results are robust to using the minimum distance to the Mexican border as an IV. Since the Gulf of Mexico separates Texas from Florida, we assign Monroe County, FL as being on the southern border. We then draw a line between the centroid of Cameron County, TX and Monroe County, FL and record the latitude-longitude coordinates along this new border at 20 mile intervals. We treat these coordinates as the centroids of "pseudo-counties" along the southern border that crosses the gulf. In untabulated results, we see that all of the findings of the paper are robust to using the minimum distance to the Mexican border as an alternative IV.

IA3 Additional Measures of Financial Sophistication

During the account opening process, retail investors are asked to report their degree of financial knowledge and whether they have gained financial experience by investing at other brokerages. We further control for retail investors' financial sophistication by adding these self-reported financial knowledge and investment experience variables to the main OLS and IV specifications.[IA2] The results, which can be found in the Internet Appendix (Tables IA1, IA2, IA3, and IA4), suggest that our findings are robust to the inclusion of these two financial sophistication variables.

IA4 Alternative Measure: Hispanics to Total Population

Our main measure of interest is H/W, which is defined as the census count of Hispanics divided by the census count of White individuals, and measures the concentration of Hispanics relative to the White population. The choice of scaling the variable by the population of Whites is supported by the idea that we want to compare how Hispanics invest relative to their White American counterparts. A potential concern could be that localities with a small population of Whites could be misclassified as areas with a high-Hispanic concentration. We thus examine whether our stock preference results are robust to scaling the Hispanic population by the total

[IA2] The financial knowledge and investment experience variables are reported on a zero-to-four scale.

population of an area. The results, which can be found in the Internet Appendix (Tables IA5, IA6, IA7, IA8, and IA9), suggest that our findings are robust to scaling our main independent variable by an area's total population. We continue to observe a strong association between the Hispanic culture and overweighting of local, lottery, and high-momentum firms by both institutional and retail investors.

Figure IA1
Cumulative Real Estate Returns by H/W Quintile

This figure plots cumulative returns for the Federal Housing Finance Agency's purchase only house price indexes for the 100 largest CBSAs in the country. These CBSAs are sorted on their concentration of Hispanic residents (as measured by H/W) and the cumulative returns for these 5 categories from 1991Q1 to 2014Q3 are plotted. Quintile 1 corresponds to the lowest concentration of Hispanic residents, while Quintile 5 corresponds to the highest concentration of Hispanic residents.

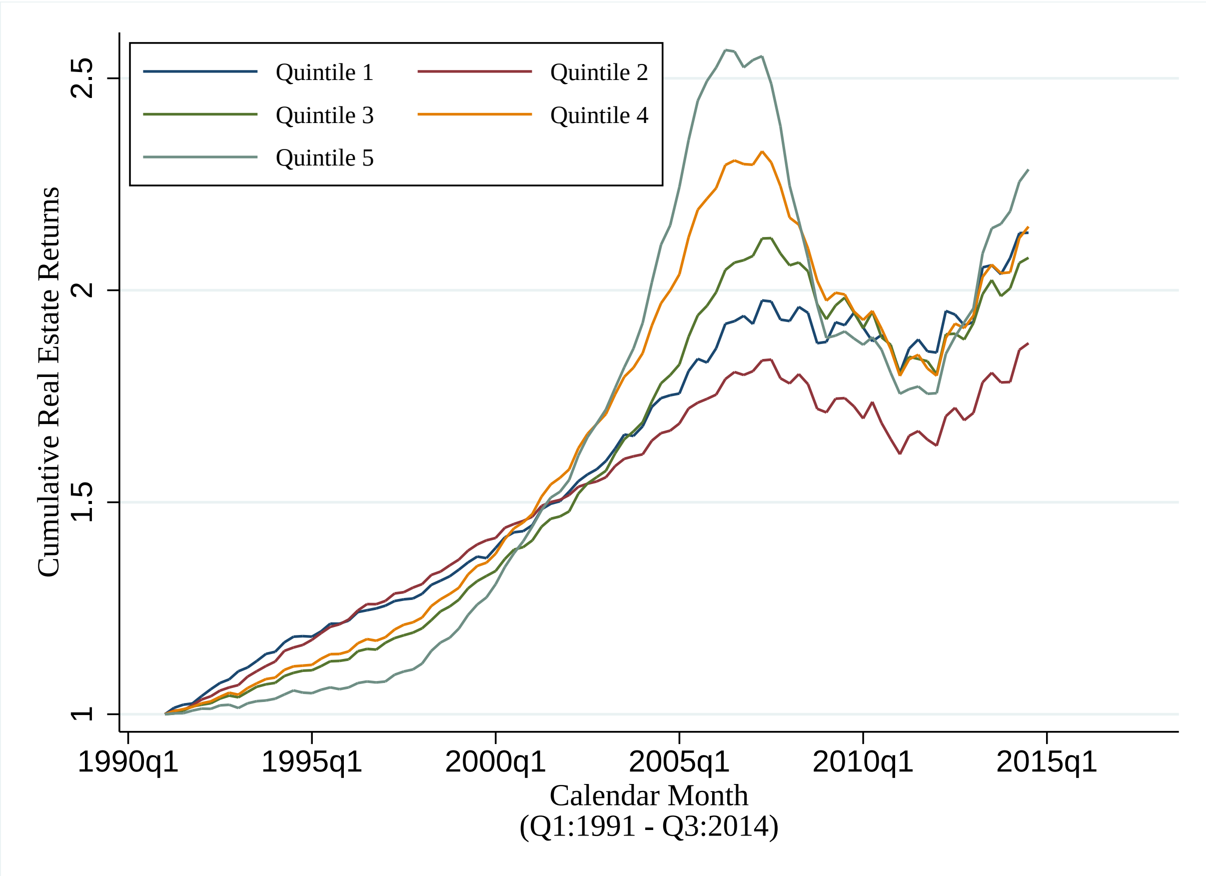

Figure IA2

Raw Real Estate Returns by H/W Quintile

This figure plots raw returns for the Federal Housing Finance Agency's purchase only house price indexes for the 100 largest CBSAs in the country. These CBSAs are sorted on their concentration of Hispanic residents (as measured by H/W) and the raw returns for these 5 categories from 1991Q1 to 2014Q3 are plotted. Quintile 1 corresponds to the lowest concentration of Hispanic residents, while Quintile 5 corresponds to the highest concentration of Hispanic residents.

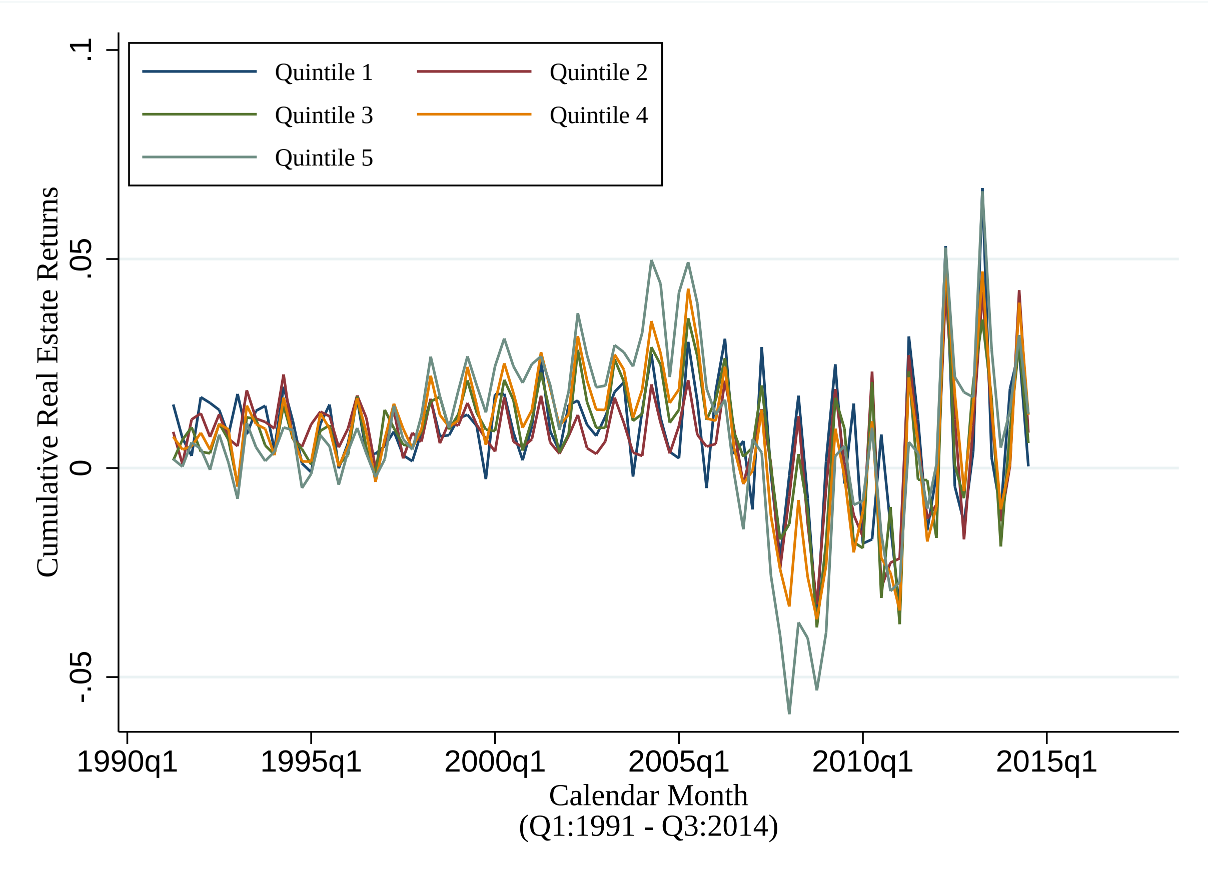

Table IA1
Hispanic Culture, Investment Preferences, and Investor Sophistication

This table presents estimates from regressions of the excess weight of household i's portfolio on the set of stocks s at time t on a vector of ZIP code, household and stock level covariates. Panel A shows results for the H/W concentration variable while Panel B uses a ZIP code's minimum distance to the Canadian border, MinDist, as an IV for the Hispanic concentration of a ZIP code. In the first two columns, s is the set of local stocks or the set of stocks headquartered within 60 miles of household i's ZIP code. In the next two columns, s is the set of lottery stocks. They are defined as firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. In columns 5-7, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. Columns (1) and (3) include ZIP code-level explanatory variables and household level regressors while columns (2) and (4) - (7) add MSA fixed effects to the model. H/W and B/W are the ratios of Hispanic and Black, respectively, to White individuals in household i's ZIP code. Persons is the census count of individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Knowledge and Experience are self-reported variables that measure a household's financial knowledge and investment experience on a zero-to-four scale. Other controls are included but suppressed for brevity. Density is a measure of population density or the total population of the ZIP code divided by its land area. We include several household level controls, including indicators if the head of the household is male (I_{Male}), if it's a married household (I_{Married}), and if the household owns its home ($I_{\text{Own Home}}$). We also control for household i's income category, the age of the household, and the level of education. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). The set of covariates is constant across columns (4) - (7) and all continuous regressors are standardized. Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level.

Panel A: OLS Regressions

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{\text{p} > 90}$	
	1	2	3	4	12 mo.	6 mo.
H/W	272.390	-21.342	13.814	7.557	2.542	2.168
	(8.34)	(-0.78)	(3.33)	(1.66)	(2.50)	(1.97)
Persons	7.967	41.929	-8.851	-3.759	1.161	1.778
	(0.33)	(2.16)	(-2.90)	(-1.17)	(1.46)	(2.02)
B/W	-92.020	-10.826	-4.606	-3.570	-0.454	-0.649
	(-8.59)	(-2.57)	(-2.86)	(-2.20)	(-1.23)	(-1.61)
Foreign	-481.095	-16.184	3.000	-7.725	-0.910	-0.112
	(-17.99)	(-0.90)	(0.80)	(-1.64)	(-0.84)	(-0.09)
I_{Urban}	333.246	275.073	9.877	10.751	2.554	1.270
	(5.44)	(5.12)	(1.45)	(1.47)	(1.37)	(0.62)
Knowledge	-100.110	-1.002	8.289	6.926	-1.356	-1.306
	(-7.56)	(-0.09)	(2.53)	(1.98)	(-1.68)	(-1.38)
Experience	-139.929	-21.425	-21.991	-20.038	1.789	-0.222
	(-6.24)	(-1.14)	(-6.91)	(-5.75)	(2.04)	(-0.23)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
Adj. R sq.	0.007	0.102	0.010	0.016	0.012	0.008
N	1,382,301	1,339,060	1,345,477	1,290,924	1,290,924	1,290,924

Table IA1
Hispanic Culture, Investment Preferences, and Investor Sophistication (Continued...)

Panel B: Instrumental Variable Regressions

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{p > 90}$	
	1	2	3	4	12 mo.	6 mo.
H/W	3975.824	2861.137	63.884	139.523	150.189	213.076
	(11.81)	(2.74)	(3.31)	(0.96)	(3.15)	(3.46)
Persons	-121.279	-3.065	-10.607	-5.766	-1.084	-1.430
	(-3.26)	(-0.09)	(-3.37)	(-1.44)	(-0.74)	(-0.74)
B/W	-532.135	-362.235	-10.374	-19.198	-17.940	-25.627
	(-7.54)	(-2.72)	(-3.58)	(-1.10)	(-3.03)	(-3.31)
Foreign	-2579.371	-1623.678	-25.582	-81.694	-83.669	-118.330
	(-12.86)	(-2.75)	(-2.23)	(-0.99)	(-3.07)	(-3.34)
I_{Urban}	296.676	280.831	9.610	11.571	3.470	2.579
	(4.31)	(4.71)	(1.41)	(1.55)	(1.54)	(0.94)
Knowledge	-317.478	-308.320	5.440	-7.079	-17.026	-23.690
	(-6.60)	(-2.60)	(1.51)	(-0.44)	(-3.09)	(-3.28)
Experience	1001.965	905.373	-6.421	22.450	49.325	67.683
	(9.19)	(2.68)	(-0.95)	(0.48)	(3.19)	(3.39)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
N	1,382,301	1,339,060	1,345,477	1,290,924	1,290,924	1,290,924

Table IA2
Controlling for Unobserved Geographic Heterogeneity and Investor Sophistication

This table presents estimates from regressions of the excess weight ($EW_{i,s,t}$) of household i's portfolio on the set of stocks s at time t following the methodology of Rajan and Zingales (1998). Panel A shows results for the H/W concentration variable while Panel B uses a ZIP code's minimum distance to the Canadian border, MinDist, as an IV for the Hispanic concentration of a ZIP code. The regressions include MSA-level fixed effects in an effort to further control for unobserved geographic heterogeneity. In addition, we include interaction terms of H/W and the minimum distance to the Canadian border with measures of household i's portfolio concentration. CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. In columns (1) and (2), the excess weight is measured with respect to the market weight of local stocks. A firm is defined to be "local" if it is headquartered within sixty miles of household i's ZIP code. Columns (3) and (4) present results where the dependent variable is the excess weight of household i's portfolio on lottery stocks. They are defined as firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. The remaining columns present results where the dependent variable is the excess weight of household i's portfolio on stocks in the top decile of returns over the last 12 and 6 month periods, respectively. The independent variable of interest in each model is the interaction of H/W with a measure of portfolio concentration (H/W×CONC or H/W×HERF). H/W is the ratio of Hispanic to White individuals in household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, $I_{Married}$ equals one if it is a married household, and $I_{Own\ Home}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). Knowledge and Experience are self-reported variables that measure a household's financial knowledge and investment experience on a zero-to-four scale. ZIP code-level controls are included but suppressed for brevity. Persons is the census count of individuals in household i's ZIP code. B/W is the ratio of Black to White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in H/W. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

Panel A: OLS Regressions

	HQ < 60		Lottery Stocks		$Returns_{p > 90}$			
					12 mo.		6 mo.	
	1	2	3	4				
H/W x CONC	58.303		7.942		0.855		1.669	
	(3.50)		(3.43)		(1.67)		(2.98)	
H/W x HERF		621.249		147.661		11.303		18.990
		(2.58)		(5.34)		(1.69)		(2.62)
I_{Male}	223.456	233.444	110.970	110.729	14.333	13.730	20.033	19.226
	(4.85)	(4.91)	(14.47)	(14.29)	(7.57)	(7.13)	(9.75)	(9.21)
$I_{Married}$	-31.477	-35.054	-26.078	-29.215	-3.609	-3.164	-4.465	-3.857
	(-0.73)	(-0.78)	(-4.08)	(-4.47)	(-2.36)	(-2.04)	(-2.67)	(-2.27)
$I_{Own\ Home}$	215.602	195.274	-5.335	-2.043	-1.491	-1.444	-2.678	-2.786
	(3.89)	(3.37)	(-0.59)	(-0.22)	(-0.67)	(-0.64)	(-1.13)	(-1.15)
Age	48.571	51.969	-1.997	-2.068	-9.717	-10.131	-10.015	-10.540
	(1.95)	(2.01)	(-0.59)	(-0.60)	(-11.73)	(-12.04)	(-11.09)	(-11.52)
Knowledge	-1.224	-5.665	-1.036	-0.626	-1.312	-1.222	-2.783	-2.578
	(-0.03)	(-0.14)	(-0.17)	(-0.10)	(-0.76)	(-0.69)	(-1.55)	(-1.39)
Experience	-58.381	-54.880	4.582	4.459	3.033	2.763	3.438	3.115
	(-1.64)	(-1.42)	(0.78)	(0.71)	(1.79)	(1.57)	(1.94)	(1.70)
MSA Code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.102	0.105	0.016	0.017	0.012	0.013	0.008	0.008
N	1,339,060	1,242,158	1,290,924	1,218,558	1,290,924	1,218,558	1,290,924	1,218,558

Table IA2
Controlling for Unobserved Geographic Heterogeneity and Investor Sophistication (Continued...)

Panel B: Instrumental Variable Regressions

	HQ < 60		Lottery Stocks		$Returns_{p > 90}$			
	1	2	3	4	12 mo.		6 mo.	
H/W IV x CONC	405.025		40.315		0.682		3.541	
	(9.25)		(8.81)		(0.63)		(3.01)	
H/W IV x HERF		4948.487		1189.917		32.838		100.479
		(5.83)		(15.83)		(1.90)		(5.20)
I_{Male}	223.791	229.910	110.901	109.904	14.334	13.712	20.029	19.162
	(4.84)	(4.81)	(14.44)	(14.01)	(7.57)	(7.12)	(9.75)	(9.17)
$I_{Married}$	-32.283	-29.593	-26.055	-27.867	-3.610	-3.136	-4.463	-3.751
	(-0.75)	(-0.65)	(-4.07)	(-4.23)	(-2.36)	(-2.02)	(-2.66)	(-2.20)
$I_{Own\ Home}$	201.121	174.315	-6.957	-7.669	-1.482	-1.560	-2.772	-3.226
	(3.61)	(2.99)	(-0.76)	(-0.82)	(-0.67)	(-0.69)	(-1.16)	(-1.33)
Age	72.513	72.190	0.279	2.898	-9.729	-10.028	-9.883	-10.152
	(2.92)	(2.81)	(0.08)	(0.83)	(-11.73)	(-11.90)	(-10.93)	(-11.07)
Knowledge	-3.944	-3.337	-1.364	-0.299	-1.310	-1.216	-2.802	-2.553
	(-0.10)	(-0.08)	(-0.23)	(-0.05)	(-0.76)	(-0.68)	(-1.56)	(-1.37)
Experience	-44.500	-44.280	5.966	7.283	3.025	2.822	3.518	3.336
	(-1.24)	(-1.14)	(1.01)	(1.16)	(1.78)	(1.61)	(1.99)	(1.82)
MSA Code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,339,060	1,242,158	1,290,924	1,218,558	1,290,924	1,218,558	1,290,924	1,218,558

Table IA3
First Stage Regressions: Hispanic Culture, Investment Preferences, and Investor Sophistication

This table presents univariate and first stage regressions of a ZIP code's minimum distance to the Canadian border on a ZIP code's Hispanic concentration. MinDist is the minimum distance from the Canadian border to household i's ZIP code. Persons is the census count of individuals in household i's ZIP code. B/W is the ratio of Black to White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Knowledge and Experience are self-reported variables that measure a household's financial knowledge and investment experience on a zero-to-four scale. Other controls are included but suppressed for brevity. Density is a measure of population density or the total population of the ZIP code divided by its land area. We include several household level controls, including indicators if the head of the household is male (I_{Male}), if it's a married household (I_{Married}), and if the household owns its home ($I_{\text{Own Home}}$). We also control for household i's income category, the age of the household, and the level of education. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). All non-indicator variables are standardized, and ZIP code-year clustered t-statistics are presented in parentheses below point estimates.

	2-SLS First Stage Regressions											
	HQ < 60 Miles				Lottery Stocks				$\text{Returns}_{\text{p} > 90}$			
	1	2	3	4	5	6	7	8	12 mo.	12 mo.	6 mo.	6 mo.
MinDist	0.001	0.000	0.002	0.001	0.001	0.000	0.002	0.001	0.002	0.001	0.002	0.001
	(31.16)	(24.91)	(3.86)	(4.22)	(31.89)	(25.36)	(4.57)	(4.76)	(4.57)	(4.76)	(4.57)	(4.76)
Persons		0.031		0.015		0.032		0.015		0.015		0.015
		(4.52)		(2.03)		(4.69)		(1.97)		(1.97)		(1.97)
B/W		0.121		0.122		0.118		0.119		0.119		0.119
		(9.03)		(8.94)		(9.30)		(9.20)		(9.20)		(9.20)
Foreign		0.528		0.558		0.531		0.560		0.560		0.560
		(41.01)		(33.51)		(40.84)		(33.29)		(33.29)		(33.29)
I_{Urban}		-0.009		-0.006		-0.011		-0.011		-0.011		-0.011
		(-1.09)		(-0.66)		(-1.41)		(-1.22)		(-1.22)		(-1.22)
Knowledge		0.002		0.005		0.003		0.006		0.006		0.006
		(0.53)		(1.28)		(0.76)		(1.48)		(1.48)		(1.48)
Experience		-0.007		-0.010		-0.008		-0.011		-0.011		-0.011
		(-1.75)		(-2.40)		(-1.91)		(-2.57)		(-2.57)		(-2.57)
Port. Chars	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Income	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
MSA Ind.	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.071	0.472	0.196	0.528	0.072	0.469	0.202	0.531	0.202	0.531	0.202	0.531
F-Stat.		620.69		17.77		643.09		22.67		22.67		22.67
N	1,382,301	1,382,301	1,339,060	1,339,060	1,345,477	1,345,477	1,290,924	1,290,924	1,290,924	1,290,924	1,290,924	1,290,924

Table IA4
First Stage Regressions: Controlling for Unobserved Geographic Heterogeneity and Investor Sophistication

This table presents univariate and first stage regressions of a ZIP code's minimum distance to the Canadian border interacted with a measure of portfolio concentration (H/W×CONC or H/W×HERF). CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. The independent variable of interest in each model is the interaction of MinDist with a measure of portfolio concentration (MinDist×CONC or MinDist×HERF). MinDist is the minimum distance from the Canadian border to household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, I_{Married} equals one if it is a married household, and $I_{\text{Own Home}}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). Knowledge and Experience are self-reported variables that measure a household's financial knowledge and investment experience on a zero-to-four scale. ZIP code-level controls are included but suppressed for brevity. Persons is the census count of individuals in household i's ZIP code. B/W is the ratio of Black to White individuals in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in MinDist. All specifications include MSA-level fixed effects. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

	2-SLS First Stage Regressions							
	HQ <60 Miles				Lottery Stocks			
	1	2	3	4	5	6	7	8
MinDist x CONC	0.017	0.017			0.017	0.017		
	(41.49)	(37.71)			(41.70)	(37.98)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(49.47)	(50.39)			(49.29)	(50.39)
I_{Male}		0.008		0.002		0.012		0.002
		(0.76)		(1.68)		(1.16)		(1.82)
I_{Married}		-0.007		-0.002		-0.011		-0.002
		(-0.78)		(-1.93)		(-1.19)		(-2.17)
$I_{\text{Own Home}}$		0.049		0.005		0.056		0.006
		(3.47)		(3.73)		(3.90)		(4.15)
Age		-0.026		-0.002		-0.028		-0.002
		(-5.15)		(-3.65)		(-5.41)		(-3.83)
Knowledge		0.013		0.000		0.015		0.001
		(1.62)		(0.73)		(1.84)		(1.16)
Experience		-0.017		-0.001		-0.019		-0.001
		(-2.11)		(-1.85)		(-2.29)		(-2.26)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.237	0.424	0.218	0.442	0.238	0.423	0.220	0.441
F-Stat.		1422.35		2539.55		1442.75		2539.33
N	1,339,060	1,339,060	1,242,158	1,242,158	1,290,924	1,290,924	1,218,558	1,218,558

Table IA4
First Stage Regressions: Controlling for Unobserved Geographic Heterogeneity and Investor Sophistication (Continued...)

	2-SLS First Stage Regressions							
	$\text{Returns}_{\text{p > 90}}$							
	12 mo.	12 mo.	12 mo.	12 mo.	6 mo.	6 mo.	6 mo.	6 mo.
MinDist x CONC	0.017	0.017			0.017	0.017		
	(41.70)	(37.98)			(41.70)	(37.98)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(49.29)	(50.39)			(49.29)	(50.39)
I_{Male}		0.012		0.002		0.012		0.002
		(1.16)		(1.82)		(1.16)		(1.82)
I_{Married}		-0.011		-0.002		-0.011		-0.002
		(-1.19)		(-2.17)		(-1.19)		(-2.17)
$I_{\text{Own Home}}$		0.056		0.006		0.056		0.006
		(3.90)		(4.15)		(3.90)		(4.15)
Age		-0.028		-0.002		-0.028		-0.002
		(-5.41)		(-3.83)		(-5.41)		(-3.83)
Knowledge		0.015		0.001		0.015		0.001
		(1.84)		(1.16)		(1.84)		(1.16)
Experience		-0.019		-0.001		-0.019		-0.001
		(-2.29)		(-2.26)		(-2.29)		(-2.26)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.238	0.423	0.220	0.441	0.238	0.423	0.220	0.441
F-Stat.		1442.75		2539.33		1442.75		2539.33
N	1,290,924	1,290,924	1,218,558	1,218,558	1,290,924	1,290,924	1,218,558	1,218,558

Table IA5

Alternative Coefficient for the OLS Regressions: Hispanic Culture and Investment Preferences

This table presents estimates from ordinary least squares regressions of the excess weight of investor i's portfolio on the set of stocks s at time t on a vector of local area and portfolio level covariates. The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include H/P, Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state and time fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. Columns (1) and (3) include ZIP code-level explanatory variables and household level regressors while columns (2) and (4) - (7) add MSA fixed effects to the model. We control for H/P, Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/P, Foreign, Density, $\text{I}_{\text{Married}}$, $\text{I}_{\text{Own Home}}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). The set of covariates is constant across columns (4) - (7). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors

	HQ < 60 Miles	Lottery Stocks	$\text{Returns}_{\text{p > 90}}$	
	(1)	(2)	12 mo.	6 mo.
H/P	27.619	0.034	0.103	0.066
	(4.76)	(2.44)	(6.60)	(5.07)
Urban	-49.591	0.055	0.089	0.056
	(-7.82)	(5.79)	(10.53)	(6.85)
Male/Female Ratio	-18.412	-0.014	-0.045	-0.041
	(-4.57)	(-0.79)	(-2.10)	(-2.45)
Median Age	-2.152	-0.022	-0.029	-0.038
	(-0.59)	(-1.81)	(-2.26)	(-3.37)
Port. Controls	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
State Ind.	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Adj. R sq.	0.176	0.068	0.065	0.048
N	150,213	151,027	151,027	151,027

Panel B: Retail Investors

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{\text{p > 90}}$	
	1	2	3	4	12 mo.	6 mo.
H/P	256.288	-0.879	13.023	-0.888	0.751	0.031
	(6.56)	(-0.03)	(3.71)	(-0.21)	(0.72)	(0.03)
I_{Urban}	300.984	231.090	8.956	4.235	2.527	1.436
	(5.06)	(4.47)	(1.35)	(0.59)	(1.41)	(0.73)
I_{Male}	140.978	234.192	106.353	105.846	13.803	19.412
	(2.97)	(5.34)	(14.79)	(14.36)	(7.61)	(9.88)
Age	45.107	24.449	-3.644	-6.068	-10.648	-11.184
	(1.75)	(1.03)	(-1.18)	(-1.92)	(-13.73)	(-13.17)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
Adj. R sq.	0.007	0.101	0.010	0.015	0.011	0.008
N	1,508,936	1,460,572	1,470,297	1,409,559	1,409,559	1,409,559

Table IA6

Alternative Coefficient for the IV Regressions: Hispanic Culture and Investment Preferences

This table presents estimates from instrumental variable regressions of the excess weight of investor i's portfolio on the set of stocks s at time t on a vector of local area and portfolio level covariates. The instrumental variable, MinDist, is the minimum distance from the Canadian border to investor i's locality (i.e., county or ZIP code). The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. Columns (1) and (3) include ZIP code-level explanatory variables and household level regressors while columns (2) and (4) - (7) add MSA fixed effects to the model. We control for Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/P, Foreign, Density, I_{Married}, $I_{\text{Own Home}}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). The set of covariates is constant across columns (4) - (7). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors				
	HQ < 60 Miles	Lottery Stocks	$\text{Returns}_{\text{p} > 90}$	
	(1)	(2)	12 mo.	6 mo.
H/P IV	34.595	0.202	0.151	0.246
	(2.24)	(3.62)	(2.31)	(4.15)
Urban	-54.303	0.130	0.144	0.118
	(-8.60)	(13.41)	(14.88)	(13.20)
Male/Female Ratio	-21.555	0.108	0.012	0.038
	(-4.38)	(5.82)	(0.57)	(2.07)
Median Age	-0.703	0.075	0.007	0.019
	(-0.19)	(6.48)	(0.57)	(1.76)
Port. Controls	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
State Ind.	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	No	No
N	150,217	151,031	151,031	151,031

Panel B: Retail Investors						
	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{\text{p} > 90}$	
	1	2	3	4	12 mo.	6 mo.
H/P IV	2598.973	1402.905	55.099	145.112	78.058	112.505
	(10.42)	(3.60)	(4.33)	(2.04)	(3.91)	(4.49)
IUrban	293.426	266.868	9.191	8.840	4.965	4.983
	(4.54)	(5.02)	(1.39)	(1.17)	(2.45)	(2.11)
IMale	109.208	205.311	105.801	102.629	12.100	16.933
	(2.17)	(4.56)	(14.70)	(13.49)	(6.39)	(8.06)
Age	65.842	42.874	-3.295	-3.961	-9.532	-9.561
	(2.44)	(1.75)	(-1.06)	(-1.18)	(-11.34)	(-10.06)
Port. Chars	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes	Yes	Yes
MSA Ind.	No	Yes	No	Yes	Yes	Yes
N	1,508,936	1,460,572	1,470,297	1,409,559	1,409,559	1,409,559

Table IA7
Alternative Coefficient: Controlling for Unobserved Geographic Heterogeneity

This table presents estimates from regressions of the excess weight ($EW_{i,s,t}$) of household i's portfolio on the set of stocks s at time t following the methodology of Rajan and Zingales (1998). Panel A shows results for the H/P concentration variable while Panel B uses a ZIP code's minimum distance to the Canadian border, MinDist, as an IV for the Hispanic concentration of a ZIP code. The regressions include MSA-level fixed effects in an effort to further control for unobserved geographic heterogeneity. In addition, we include interaction terms of H/P and the minimum distance to the Canadian border with measures of household i's portfolio concentration. CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. In columns (1) and (2), the excess weight is measured with respect to the market weight of local stocks. A firm is defined to be "local" if it is headquartered within sixty miles of household i's ZIP code. Columns (3) and (4) present results where the dependent variable is the excess weight of household i's portfolio on lottery stocks. They are defined as firms in the lowest 50^{th} stock price percentile, the highest 50^{th} idiosyncratic volatility percentile, and the highest 50^{th} skewness percentile. All three sorts are carried out independently. The remaining columns present results where the dependent variable is the excess weight of household i's portfolio on stocks in the top decile of returns over the last 12 and 6 month periods, respectively. The independent variable of interest in each model is the interaction of H/P with a measure of portfolio concentration (H/P×CONC or H/P×HERF). H/P is the ratio of Hispanics to the total population in household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, $I_{Married}$ equals one if it is a married household, and $I_{Own\ Home}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). ZIP code-level controls are included but suppressed for brevity. B/P is the ratio of Blacks to the total population in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in H/P. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

Panel A: OLS Regressions

	HQ < 60		Lottery Stocks		$Returns_{p > 90}$			
					12 mo.		6 mo.	
	1	2	3	4				
H/P x CONC	169.499		19.555		1.765		3.909	
	(6.18)		(5.42)		(2.10)		(4.12)	
H/P x HERF		2326.270		467.347		28.492		56.585
		(5.07)		(8.65)		(2.19)		(3.84)
I_{Male}	236.657	247.816	105.896	105.411	13.839	13.086	19.458	18.491
	(5.40)	(5.48)	(14.38)	(14.14)	(7.63)	(7.09)	(9.90)	(9.25)
$I_{Married}$	-29.162	-34.556	-22.266	-25.021	-3.822	-3.336	-4.450	-3.817
	(-0.71)	(-0.80)	(-3.65)	(-4.00)	(-2.63)	(-2.25)	(-2.79)	(-2.35)
$I_{Own\ Home}$	217.261	197.674	-9.396	-6.225	-2.135	-2.374	-3.419	-3.748
	(4.14)	(3.61)	(-1.07)	(-0.70)	(-1.01)	(-1.10)	(-1.49)	(-1.61)
Age	31.811	41.430	-5.139	-4.425	-10.583	-10.859	-11.018	-11.374
	(1.33)	(1.69)	(-1.63)	(-1.38)	(-13.63)	(-13.75)	(-12.97)	(-13.21)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.101	0.104	0.015	0.016	0.011	0.012	0.008	0.008
N	1,460,572	1,354,339	1,409,559	1,329,347	1,409,559	1,329,347	1,409,559	1,329,347

Table IA7

Alternative Coefficient: Controlling for Unobserved Geographic Heterogeneity (Continued...)

Panel B: Instrumental Variable Regressions								
	HQ < 60		Lottery Stocks		$\text{Returns}_{p > 90}$			
	1	2	3	4	12 mo.		6 mo.	
H/P IV x CONC	568.096		65.939		3.011		7.864	
	(9.98)		(10.91)		(2.07)		(4.98)	
H/P IV x HERF		7393.059		1952.665		71.431		186.176
		(6.44)		(18.65)		(2.92)		(6.81)
I_{Male}	239.975	251.548	106.216	106.437	13.848	13.116	19.485	18.580
	(5.46)	(5.54)	(14.42)	(14.20)	(7.64)	(7.10)	(9.92)	(9.29)
I_{Married}	-27.592	-31.643	-22.086	-24.185	-3.817	-3.311	-4.435	-3.744
	(-0.67)	(-0.74)	(-3.61)	(-3.85)	(-2.62)	(-2.24)	(-2.78)	(-2.30)
$I_{\text{Own Home}}$	208.168	185.551	-10.521	-10.086	-2.165	-2.486	-3.515	-4.085
	(3.97)	(3.38)	(-1.20)	(-1.13)	(-1.02)	(-1.15)	(-1.54)	(-1.75)
Age	49.911	56.441	-3.047	0.000	-10.526	-10.731	-10.840	-10.988
	(2.10)	(2.32)	(-0.96)	(0.00)	(-13.55)	(-13.59)	(-12.75)	(-12.75)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP Code Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,460,572	1,354,339	1,409,559	1,329,347	1,409,559	1,329,347	1,409,559	1,329,347

Table IA8
First Stage Regressions for the Alternative Coefficient: Hispanic Culture and Investment Preferences

This table presents univariate and first stage regressions of an area's minimum distance to the Canadian border on the local Hispanic concentration measure. The instrumental variable, MinDist, is the minimum distance from the Canadian border to investor i's locality (i.e., county or ZIP code). The institutional-level analysis is at the county level and the retail investor-level analysis is at the ZIP code level. For the first dependent variable, s is the set of local stocks or the set of stocks headquartered within 60 miles of investor i's county/ZIP code. For the second dependent variable, s is the set of lottery stocks. They are defined as firms in the highest 50^{th} idiosyncratic volatility percentile and the highest 50^{th} skewness percentile. For the retail investor sample, we include a third condition: firms in the lowest 50^{th} stock price percentile. The sorts are carried out independently. In the last columns, the set of stocks s includes those in the top decile of returns over the most recent 12 and 6 month periods, respectively. In Panel A, the sample is the quarterly stock holdings of 13(f) institutions compiled by Thomson Reuters from 1980 to 2018. It excludes banks and insurance companies. The controls include Persons, Urban, Male/Female Ratio, and Median Age. We also include the following controls: Portfolio Value, HERF, Education, Married. These are excluded for brevity. The specifications include state fixed effects. All variables are standardized and county-year-quarter clustered t-statistics are presented in parentheses below point estimates. The sample in Panel B is from a discount brokerage account from January 1991 to November 1996. We control for Persons, I_{Urban}, I_{Male}, and Age. Other controls are included but suppressed for brevity. We also include B/P, Foreign, Density, I_{Married}, $I_{\text{Own Home}}$, Income, Level of Education, Sharpe Ratio (SR), Alpha, and average Portfolio Return, and the total Portfolio Value (in thousands). Heteroskedasticity robust t-statistics are presented in parentheses and are clustered at the ZIP code-year level. All continuous regressors are standardized. The variables are defined in Appendix Table A1.

Panel A: Institutional Investors

	HQ < 60 Miles		Lottery Stocks		$\text{Returns}_{p > 90}$			
	(1)	(2)	(3)	(4)	12 mo.		6 mo.	
MinDist	0.213	0.080	0.000	0.001	0.000	0.001	0.000	0.001
	(5.80)	(1.94)	(1.03)	(4.95)	(2.88)	(3.29)	(2.94)	(5.34)
Urban		-49.630		0.056		0.089		0.056
		(-7.83)		(5.91)		(10.63)		(7.00)
Male/Female Ratio		-17.756		-0.014		-0.043		-0.041
		(-4.30)		(-0.81)		(-2.03)		(-2.44)
Median Age		-2.691		-0.014		-0.028		-0.032
		(-0.80)		(-1.20)		(-2.17)		(-2.89)
Port. Controls	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	No	Yes	No	Yes	No	Yes	No	Yes
State Ind.	No	Yes	No	Yes	No	Yes	No	Yes
Year-Quarter FE	No	No	No	No	No	No	No	No
Adj. R sq.	0.010	0.176	0.000	0.068	0.000	0.065	0.000	0.048
F-Stat.		1326.38		1309.17		1309.17		1309.17
N	150,213	150,213	151,027	151,027	151,027	151,027	151,027	151,027

Table IA8
First Stage Regressions for the Alternative Coefficient: Hispanic Culture and Investment Preferences (Continued...)

Panel B: Retail Investors												
	HQ < 60 Miles				Lottery Stocks				$\text{Returns}_{\text{p} > 90}$			
	1	2	3	4	5	6	7	8	12 mo.	12 mo.	6 mo.	6 mo.
Min. Dist.	0.001	0.001	0.003	0.002	0.001	0.001	0.003	0.002	0.003	0.002	0.003	0.002
	(27.04)	(20.04)	(6.85)	(7.46)	(27.67)	(20.48)	(7.21)	(7.57)	(7.21)	(7.57)	(7.21)	(7.57)
I_{Urban}		0.568		0.547		0.572		0.548		0.548		0.548
		(41.58)		(33.03)		(41.55)		(32.86)		(32.86)		(32.86)
I_{Male}		0.019		0.021		0.018		0.022		0.022		0.022
		(2.93)		(3.74)		(2.74)		(4.02)		(4.02)		(4.02)
Age		-0.014		-0.012		-0.014		-0.013		-0.013		-0.013
		(-4.86)		(-4.57)		(-4.71)		(-5.14)		(-5.14)		(-5.14)
Port. Chars	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Income	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
MSA Ind.	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R sq.	0.115	0.509	0.349	0.640	0.115	0.498	0.366	0.648	0.366	0.648	0.366	0.648
F-Stat.		401.73		55.61		419.23		57.30		57.30		57.30
N	1,508,936	1,508,936	1,460,572	1,460,572	1,470,297	1,470,297	1,409,559	1,409,559	1,409,559	1,409,559	1,409,559	1,409,559

Table IA9
First Stage Regressions for the Alternative Coefficient: Controlling for Unobserved Geographic Heterogeneity

This table presents univariate and first stage regressions of a ZIP code's minimum distance to the Canadian border interacted with a measure of portfolio concentration (H/P×CONC or H/P×HERF). CONC is measured as 10 minus the number of stocks in household i's portfolio. HERF is a Herfindahl Index of household i's portfolio if i is a long-only investor. The independent variable of interest in each model is the interaction of MinDist with a measure of portfolio concentration (MinDist×CONC or MinDist×HERF). MinDist is the minimum distance from the Canadian border to household i's ZIP code. Controls for the household's demographics and portfolio characteristics are included. I_{Male} equals one if the head of the household is male, I_{Married} equals one if it is a married household, and $I_{\text{Own Home}}$ equals one if the household owns its home. We also control for household i's income category and the age of the household. Additional portfolio controls include the Sharpe Ratio (SR), Alpha, and average Portfolio Return of the household portfolio over the sample period and the total Portfolio Value (in thousands). ZIP code-level controls are included but suppressed for brevity. B/P is the ratio of Blacks to the total population in household i's ZIP code. Foreign is the proportion of foreign born individuals in the ZIP code. I_{Urban} is an indicator variable taking on a value of one if the Census Bureau classifies household i's ZIP code as urban. Density is a measure of population density or the total population of the ZIP code divided by its land area. Note that interaction terms are not standardized to aid in comparisons across differences in MinDist. All specifications include MSA-level fixed effects. Standard errors are clustered at the ZIP code-year level and are included in parentheses below point estimates.

2-SLS First Stage Regressions								
	HQ < 60 Miles				Lottery Stocks			
	1	2	3	4	5	6	7	8
MinDist x CONC	0.011	0.011			0.011	0.011		
	(41.41)	(38.83)			(41.13)	(38.64)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(52.91)	(52.96)			(53.14)	(53.27)
I_{Male}		0.002		0.000		0.004		0.000
		(0.34)		(0.62)		(0.86)		(1.02)
I_{Married}		-0.008		-0.001		-0.009		-0.001
		(-1.91)		(-1.92)		(-2.07)		(-2.20)
$I_{\text{Own Home}}$		0.029		0.003		0.031		0.003
		(4.37)		(4.69)		(4.55)		(5.15)
Age		-0.014		-0.001		-0.014		-0.001
		(-6.28)		(-5.03)		(-6.13)		(-5.15)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.381	0.517	0.375	0.567	0.380	0.512	0.378	0.566
F-Stat.		1508.06		2804.80		1492.68		2838.12
N	1,460,572	1,460,572	1,354,339	1,354,339	1,409,559	1,409,559	1,329,347	1,329,347

Table IA9

First Stage Regressions for the Alternative Coefficient: Controlling for Unobserved Geographic Heterogeneity (Continued...)

	2-SLS First Stage Regressions							
	$Returns_{p>90}$							
	12 mo.	12 mo.	12 mo.	12 mo.	6 mo.	6 mo.	6 mo.	6 mo.
MinDist x CONC	0.011	0.011			0.011	0.011		
	(41.13)	(38.64)			(41.13)	(38.64)		
MinDist x HERF			0.000	0.000			0.000	0.000
			(53.14)	(53.27)			(53.14)	(53.27)
I_{Male}		0.004		0.000		0.004		0.000
		(0.86)		(1.02)		(0.86)		(1.02)
$I_{Married}$		-0.009		-0.001		-0.009		-0.001
		(-2.07)		(-2.20)		(-2.07)		(-2.20)
$I_{Own\ Home}$		0.031		0.003		0.031		0.003
		(4.55)		(5.15)		(4.55)		(5.15)
Age		-0.014		-0.001		-0.014		-0.001
		(-6.13)		(-5.15)		(-6.13)		(-5.15)
MSA Ind.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	No	Yes	No	Yes	No	Yes
ZIP Code Controls	No	Yes	No	Yes	No	Yes	No	Yes
Portfolio Chars.	No	Yes	No	Yes	No	Yes	No	Yes
Adj. R sq.	0.380	0.512	0.378	0.566	0.380	0.512	0.378	0.566
F-Stat.		1492.68		2838.12		1492.68		2838.12
N	1,409,559	1,409,559	1,329,347	1,329,347	1,409,559	1,409,559	1,329,347	1,329,347