THE JOURNAL OF FINANCE • VOL. LX, NO. 1 • FEBRUARY 2005

Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments

ZORAN IVKOVIĆ and SCOTT WEISBENNER*

ABSTRACT

Using data on the investments a large number of individual investors made through a discount broker from 1991 to 1996, we find that households exhibit a strong preference for local investments. We test whether this locality bias stems from information or from simple familiarity. The average household generates an additional annualized return of 3.2% from its local holdings relative to its nonlocal holdings, suggesting that local investors can exploit local knowledge. Excess returns to investing locally are even larger among stocks not in the S&P 500 index (firms for which information asymmetries between local and nonlocal investors may be largest).

> Behold, the fool saith, "Put not all thine eggs in the one basket"—which is but a manner of saying, "Scatter your money and your attention"; but the wise man saith, "Put all your eggs in one basket and—*watch that basket.*"
>
> Mark Twain, 1894

THE FINANCE LITERATURE HAS YIELDED a large number of in-depth studies concerning the investments managed by professional money managers, yet historically, relatively little has been known about the individual investors' money management, in no small part because of the shortage of high-quality data available for academic research. This is despite the fact that United States individual investors have been holding around 50% of the stock market in direct stock investments.

The key issue we address in this paper is the availability of asymmetric information in financial markets, particularly in the context of geography of investment by individual investors. While it is well understood that there may be a number of determinants of the availability and quality of information about a company, including, for example, company size, number of analysts following

*Both authors are from the Department of Finance, University of Illinois at Urbana-Champaign. Scott Weisbenner is with the NBER. We would like to extend our gratitude to an anonymous discount broker for providing the data on individual investors' positions, trades, and demographics. Special thanks to Terry Odean for his help in obtaining and understanding the data set. We thank Alok Kumar, Allen Poteshman, Mark Seasholes, and an anonymous referee for useful insights and helpful discussions. Both authors acknowledge the financial support from the College Research Board at the University of Illinois at Urbana-Champaign. Finally, we thank seminar participants at the University of Illinois and the 2003 Western Finance Association Meetings for their comments and constructive suggestions.

it, and media coverage, it is reasonable to hypothesize that, *ceteris paribus*, investors may be able to gather value-relevant information about the companies local to them (henceforth local companies) with greater ease and accuracy than they could about remote companies (henceforth nonlocal companies). Indeed, Coval and Moskowitz (2001) demonstrate that professional managers' local investments outperform their remote investments, a finding that both provides a richer characterization of professional managers' skill and, more importantly in the context of this paper, suggests that proximity to investment opportunities facilitates the acquisition of disproportionately accurate value-relevant information.

Using a detailed data set on the investments that 78,000 U.S. retail investors made through a large discount broker over the 6-year period from 1991 to 1996, we find that individual investors exhibit local bias, that is, disproportionate preference for local stocks, to an even larger degree than U.S. mutual fund managers do (see Coval and Moskowitz (1999)). We find that the average share of local investments (defined as investments in companies headquartered within 250 miles from the investor) is around 30%, both in terms of the number of stocks in the household portfolio and their value. This figure is disproportionately high—nearly 20 percentage points higher than the average percent of all firms headquartered within 250 miles from the household (both equally and value weighted).

While it appears that the professional managers' preferences for (and especially success in) pursuing local investment opportunities stem from their ability to exploit the ensuing information asymmetries (Coval and Moskowitz (2001)), it is not clear whether the individual investors' local bias is primarily a result of their ability to exploit asymmetric information or a result of their inclination simply to invest into the companies they are familiar with (though not necessarily particularly informed about).

That households would have a pronounced preference for local stocks is not unexpected. There is evidence that people tend to invest disproportionately in the companies they are relatively familiar with, be it by virtue of their inclination to favor domestic investments over investments in foreign assets—the so-called home bias,[1] or, in the domestic arena, invest disproportionately into own-company stock (Benartzi (2001), Liang and Weisbenner (2002)). If individual investors' local investments stem primarily from noninformation-based reasons, as the evidence presented in Huberman (2001), Grinblatt and Keloharju (2001), and Zhu (2003) suggests, then investors' locality should be interpreted primarily as a behavioral phenomenon. If locality stems from households investing in the companies they are familiar with, though not necessarily particularly informed about, local investments likely will not earn superior returns on average.[2] Indeed, Benartzi finds in the context of 401(k) plans that allocations

[1] For key results related to home bias, see, for example French and Poterba (1991), Gehrig (1993), Cooper and Kaplanis (1994), and Tesar and Werner (1997).

[2] Zhu (2003) explores the local bias of individual investors using the same data set as we do and finds that local bias is driven by simple familiarity, a conclusion different from ours. The

to company stock, the most local of all investments, do not predict future company stock returns.

On the other hand, if the individual investors' local bias is information-driven, we should uncover the evidence of superiority of their local investments, at least relative to their nonlocal investments. Such a finding would shed some light not only on individual investors' ability to identify and exploit information asymmetries, but also on the interpretation of the professional managers' investment skill. Indeed, if individual investors could generate superior returns on their local investments, then such an accomplishment by professional managers would seem less impressive and should be reinterpreted primarily as the skill to realize that there is investment value in uncovering local information and to collect it, rather than the skill to carry out particularly insightful analyses of such information.

The central question we address is whether local investments are based on value-relevant information, in which case those investments should do well on average, or whether they stem primarily from households investing in the companies they are familiar with, though not necessarily particularly informed about, in which case the local investments likely will not earn superior returns on average. Thus, we next address the relation between location-driven information asymmetry and household portfolio returns. Simply put, if individual investors succeed in collecting and processing locally available value-relevant information, then the value of that information should be reflected in the performance of their portfolios.

In our initial inquiry, we find that local investments outperformed nonlocal ones by 3.2% per year over a 1-year horizon. The excess return to investing locally is even larger among stocks not in the S&P 500 index (firms for which information asymmetries between local and nonlocal investors may be largest), while no excess return is earned by households that invest in local S&P 500 stocks (for which information asymmetries are likely to be smallest). Moreover, the evidence is robust to the choice of the period of returns measurement, various adjustments for risk, industry composition of the portfolio, location of the

approach followed therein differs from ours along three important dimensions. First, our sample size is larger. Indeed, while Zhu identifies only 27,189 households with common-stock positions and zip-code information, we identify around 55,000 such households. Moreover, we identify the location of 7,353 firms that were held by at least one household, while Zhu identifies only 5,707 such firms. Second, the two papers control for the geographic distribution of households and firms differently. Our principal measure of a household's local bias is the fraction of its portfolio headquartered within 250 miles minus the fraction of the market portfolio headquartered within 250 miles. By contrast, the measure employed by Zhu is the difference between the distance between the household and the market and the distance between the household and its portfolio. Finally, in seeking to identify the source of local bias—simple familiarity versus availability of value-relevant information—Zhu first classifies household portfolios into quintiles by their local bias and then compares the performance of the five portfolios, whereas we employ a variety of approaches. The two most prominent approaches are regressing household portfolio returns on the portfolio weights of local and nonlocal investments and forming and evaluating the performance of zero-cost portfolios by aggregating all the positions local to the respective households (within a 250-mile radius) into the long side of such a portfolio and aggregating all the positions nonlocal to the respective households into its short side.

household, and an alternative explanation of insider trading on own/former employers' stock.

Having established that household portfolio performance is positively related to the extent and nature of locality in investors' portfolios and that the bulk of the performance differential stems from the investments in less widely known (non-S&P 500) stocks, we explore the benefits of pursuing locality by mimicking households' local investments and shying away from their nonlocal investments. To that end, we first dissect common stocks in each household portfolio according to two dimensions—locality of the stocks and their S&P 500 status (our proxy for information asymmetry)—and then form aggregate position-weighted portfolios across all households. The results support the hypothesis that locally available information is value-relevant. Specifically, the zero-cost portfolio that takes a long position in local investments and a short position in nonlocal investments has statistically and economically significant returns: Monthly raw (risk-adjusted) returns are 17 (9) basis points. This effect stems primarily from the stocks that exhibit more information asymmetry. Indeed, the corresponding raw returns and risk-adjusted returns are substantially smaller for S&P 500 stocks (11 and 6 basis points per month, respectively) than for non-S&P 500 stocks (25 and 20 basis points per month, respectively). These results are striking, as the only portfolio formation criteria were geography and a basic stock index membership characteristic. Even these simple breakdowns yielded abnormal performance of 2.4–3.0% per year, and adjustment for risk characteristics of the portfolios made little difference.

The final question we address is whether simply observing a firm's local ownership (i.e., the collective opinion of the potential investors local to the firm) can compensate for the effect of locally available value-relevant information. In other words, if the investor is not local, should he join the locals (i.e., follow the actions of the investors local to the firm)? The results suggest that merely favoring nonlocal non-S&P 500 stocks with high local ownership goes a long way toward reducing and almost eliminating the information asymmetry between the local and nonlocal non-S&P 500 stocks. Simply put, mimicking what locals do is about as good as being local.

The remainder of this paper is organized as follows. Section I presents the data and summary statistics. Section II considers the characteristics and determinants of locality of investment from the perspective of the household. Section III focuses on the determinants of a firm's local ownership. Section IV seeks to characterize and quantify the information asymmetry related to locality of investment and provides a number of robustness checks. Finally, Section V concludes.

I. Data and Descriptive Statistics

A. Data

We compile the data for this paper from several sources. The primary source is a large data set, obtained from a large brokerage house, of individual investors'

monthly positions and trades over a 6-year period from 1991 to 1996. The data set covers all the investments that 78,000 households made through the brokerage house, such as common stocks, mutual funds, government bonds, corporate bonds, foreign securities, and derivatives. Households in the data set have as few as one account and as many as 21 accounts (the median is 2). The data set also provides some demographic information about the households, such as their zip codes, self-reported incomes, occupations, and ages. For a detailed description of the data set, see Barber and Odean (2000).

In this paper we focus on the common stocks traded on the NYSE, AMEX, and Nasdaq exchanges. Common stock investments constitute roughly three quarters of the total value of household investments through the brokerage house in the sample. In the remainder of the paper we use the term total household portfolio as the shorthand for the total household investment through the brokerage house—each household could have investments through other brokerage houses and investment channels, about which we have no data. We use the Center for Research in Security Prices (CRSP) database to obtain information on stock prices and returns and COMPUSTAT to obtain several firm characteristics, including the location of the company headquarters (identified by its state and county codes). We use the headquarters location as opposed to the state of incorporation because firms often do not have the majority of their operations in their state of incorporation.

The sample of households used in this study is a subset of the entire collection of households for which we could ascertain their zip codes and thus determine their locations. We identify around 11,300 distinct zip codes in the data. We exclude from the sample households that were not located in the continental United States (that is, investors from Hawaii, Alaska, and Puerto Rico) and several zip codes corresponding to military installations. We obtain the latitude and longitude for each of the zip codes from the U.S. Census Bureau's Gazetteer Place and Zip Code Database. The corresponding company location comes from the COMPUSTAT Annual Research Files, which contain the information regarding the company headquarters' county code. Finally, we identify the latitude and longitude for each county from the U.S. Census Bureau's Gazetteer Place and Zip Code Database as well. We use the standard formula for computing the distance $d(a, b)$ in statutory miles between two points, a and b, as follows:

$$d(a,b) = \arccos\{\cos(a_1)\cos(a_2)\cos(b_1)\cos(b_2) + \cos(a_1)\sin(a_2)\cos(b_1)\sin(b_2) + \sin(a_1)\sin(b_1)\}r, \quad (1)$$

where a_1 and b_1 (a_2 and b_2) are the latitudes (longitudes) of the two points (expressed in radians), respectively, and r denotes the radius of the Earth (approximately 3,963 statutory miles).

We exclude the stocks we could not match with CRSP and COMPUSTAT; they were most likely listed on smaller exchanges. We also exclude stocks not headquartered in the continental United States. The resulting "market"—the universe of stocks that we could obtain the necessary characteristics and

information about—consists of 5,478 stocks at the end of 1991, which covers 89% of the overall market capitalization at the end of 1991.

In each sample year, we discard the households with a total portfolio value under $1,000 as of the end of the previous year. In sum, the resulting sample of households holding at least one common stock that we could identify and characterize in at least one of their account portfolios consists of 34,517 households at the end of 1991. This number tends to decrease over the sample period, primarily reflecting the liquidation of some accounts (the trades in the last month sell off the positions at the end of the previous month and there are no subsequent positions or trades recorded for the account).

B. Descriptive Statistics

The basic household sample characteristics are presented in Table I. The table reports the summary statistics of household income and household portfolio statistics at the end of 1991.[3] Panel A presents the income statistics. Households could report annual income in nine ranges (0–15, 15–20, 20–30, 30–40, 40–50, 50–75, 75–100, 100–125, and 125+ thousands of dollars). One-eighth of the households (12%) did not provide income information. We assume that the income level for the households in the first eight categories was the midpoint of the income range. Based on the calculations using the 1992 Survey of Consumer Finances (SCF), the average household income in 1991 for households that had at least $125,000 of income was $250,000 (see Kennickell and Starr-McCluer (1994) for a description of the SCF data set). Thus, we assign the income level of $250,000 to the households that reported income >$125,000. The median household that reported its annual income reported it in the range from $50,000 to $75,000 (the midpoint presented in the table is $63,000) and the interquartile range of income range midpoints is $45,000–113,000. Therefore, households in the sample are clearly more affluent than the average household in the United States.

Panel B focuses on household portfolios of NYSE, AMEX, and Nasdaq common stocks as documented by the position files at the end of 1991. The mean value of common stocks in household portfolios was $32,911, while the median was $12,416. The average household held relatively few stocks (on average 3.3 stocks; the median was only two stocks per household). Around 38% of the households in the sample held assets other than common stocks (most often mutual funds). Among such households, around one-third of the average household portfolio was allocated to assets other than common stocks (36% average; 30% median). Thus, the common-stock portions of household portfolios in the sample typically consisted of a handful of stocks, which precludes the possibility of extensive diversification among common stocks.[4] We focus on these three

[3] Portfolio summary statistics for subsequent years (not reported) are qualitatively similar to those presented in Table I.

[4] While diversification could be more pronounced for the households holding other asset classes, the summary statistics suggest that nearly two-thirds of all households in the sample (62%)

Table I

Summary Statistics of Household Income and 1991 Household Portfolios

The sample consists of 34,517 households that held at least $1,000 of stocks in their portfolios at the end of 1991. Panel A presents the income statistics. Households could report annual income in nine ranges (0–15, 15–20, 20–30, 30–40, 40–50, 50–75, 75–100, 100–125, and 125+ thousands of dollars). One-eighth of the households did not provide income information. The income level for the households in the first eight categories is assumed to be the midpoint of the income range. Households that reported income greater than $125,000 are assigned an income level of $250,000 because calculations using the 1992 Survey of Consumer Finances suggest that the average household income in 1991 for households that had at least $125,000 of income was $250,000. Panel B presents the basic household portfolio statistics. Panel C presents the household portfolio statistics related to locality of households' investments. Household portfolio distance measures are computed by value-weighting the distance measures across individual stocks in the portfolio according to their household portfolio equity position.

	Mean	Median	25–75th Percentile
Panel A: Income Statistics			
Income < $50,000	25%		
Income between $50,000 and $100,000	39%		
Income over $100,000	23%		
Household did not report income	12%		
Income level in thousands (conditional on reporting a range)	$89	$63	$45–113
Panel B: Basic Portfolio Statistics			
Stock portfolio value	$32,911	$12,416	$6,225–27,938
Number of stocks in portfolio	3.3	2	1–4
Household holds assets other than common stocks	38%		
Ratio of other assets to total assets in portfolio (conditional on holding other assets)	36%	30%	11–57%
Panel C: Distance Portfolio Statistics			
Distance to portfolio (miles)	917	796	311–1,391
Distance to market portfolio	1,225	1,030	768–1,810
Distance to portfolio–distance to market portfolio	−308	−251	−705–141
Percent of portfolio within 250 miles	31.5	2.8	0.0–65.6
Percent of all firms within 250 miles	12.6	5.7	4.3–15.1
Percent of portfolio within 250 miles–percent of all firms within 250 miles	18.9	−0.8	−5.7–48.6

household portfolio characteristics and on household income because we view them as (admittedly noisy) measures of two important facets of investment:

typically held only few common stocks and were thus largely undiversified. Moreover, Goetzmann and Kumar (2001) show that the investors' portfolios were underdiversified relative to the extent of diversification possible when holding few stocks.

propensity toward diversification (larger number of stocks, presence of asset classes other than common stocks) and availability of resources to collect information (larger dollar value of the portfolio and household income).

Table I (panel C) further summarizes the household portfolio statistics pertaining to the locality of their common stock investments according to two metrics. The first is a simple comparison of the distance from the household to its portfolio and the distance from the household to the market portfolio (the upper three rows in the panel). The second metric employs household proximity to company headquarters, where we set the threshold of locality at the distance of 250 miles (the lower three rows in the panel). Household portfolio distance measures were computed by value-weighting the distance measures across individual stocks in the portfolio according to their household portfolio equity position.

Table I (panel C) also provides evidence of considerable local bias. The average portfolio is closer to the household than the market is by about 300 miles, although the nontrivial interquartile range suggests a broad range of locality, from households substantially closer to their portfolio than to the market (by around 700 miles or more) to those that are somewhat further away from their portfolio than from the market (by around 140 miles or more).

In most subsequent analyses, rather than calculating the average distance to investments, we calculate the proportion of the investments that are geographically close to the household. We regard 250 miles as a plausible upper bound on the span of local information and regard all investments in companies whose headquarters are within that distance as local. The distance of 250 miles is reachable with a daily round trip by car; local newspapers, radio, TV, and other media occasionally provide some coverage of the local events within that distance. This cutoff is admittedly arbitrary and it could be argued that it is on the conservative side.[5] Nevertheless, we identify strong locality effects that match economic intuition based on information asymmetry.

In the sample of households at the end of 1991 the mean (median) percentage of the number of stocks in the household portfolio headquartered within 250 miles from the investor is 31.5% (2.8%). In fact, slightly more than one-half of the households in the sample (17,727 of 34,517) invest in at least one local firm and 30% of the households tilt their portfolio more toward local than toward nonlocal stocks, which could be driven in part by the geographic distribution of investors and companies (as shown in Figure 1, New York and California boast disproportionately high numbers of both).

We control for the geographic distribution of companies relative to investors by subtracting from this percentage another percentage—the fraction of all firms within 250 miles (presented in the fifth row of Table I, panel C). The resulting measure of household locality is summarized in the last row of Table I (panel C). It is similar to the simple distribution of the percentage of portfolio

[5] As a robustness check, we replicate the key analyses with the locality cutoff set at 100 km (see Section IV.H). The results are even stronger, which indeed suggests, that setting the cutoff at 250 miles is conservative.

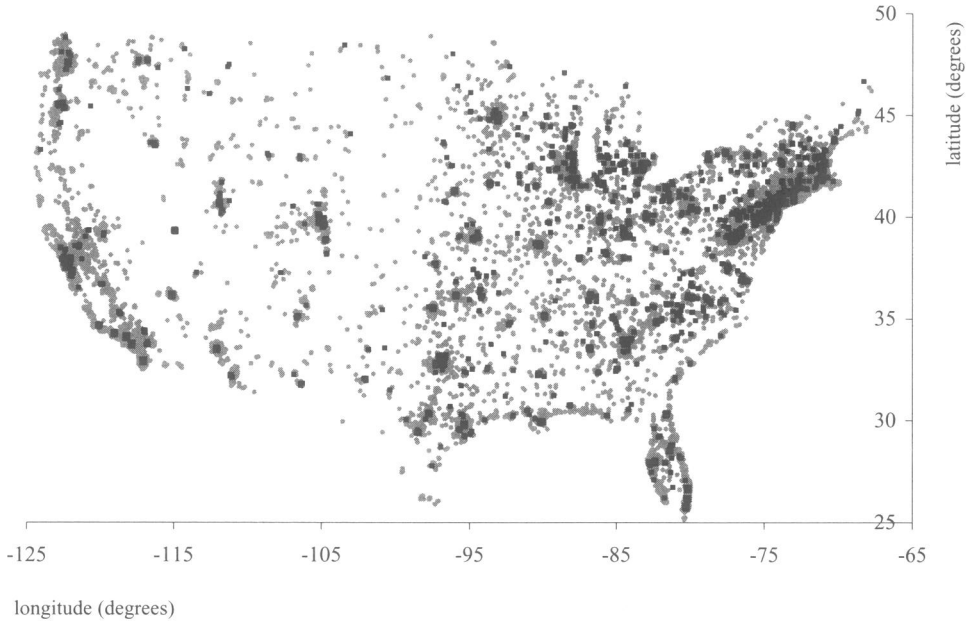

Figure 1. Distribution of U.S. publicly traded firms (black) and retail investors (gray), 1991. The figure presents the geographic distribution of 34,517 households from the sample and of 5,478 company headquarters, both located in the continental United States, at the end of 1991. The sample of households depicted in this figure is a subset of the entire collection of households that had a total portfolio of at least $1,000 at the end of 1991 and for which we could ascertain their zip code and thus determine their location. We identify around 11,300 distinct zip codes in the data. We exclude from the sample households that were not located in the continental United States (i.e., investors from Hawaii, Alaska, and Puerto Rico) and several zip codes corresponding to military installations. We obtain the latitude and longitude for each of the zip codes from the U.S. Census Bureau's Gazetteer Place and Zip Code Database. The corresponding company locations come from COMPUSTAT Annual Research Files, which contain the information regarding the company headquarters' county code. Finally, we identify the latitude and longitude for each county from the U.S. Census Bureau's Gazetteer Place and Zip Code Database as well. We exclude the stocks that we could not match with CRSP and COMPUSTAT, as well as the stocks not headquartered in the continental United States. The resulting "market"—the universe of stocks that we could obtain the necessary characteristics and information about—depicted in the figure consists of 5,478 stocks at the end of 1991, which cover 89% of the overall market capitalization at the end of 1991.

stocks within 250 miles. The mean percentage of local firms that households invest in exceeds the percentage of all local firms by 18.9% on average. It ranges from −41% to 99.95% at the end of 1991. The first (low) extreme is typical of the households that chose to invest only nonlocally, yet live in the area with many local investment opportunities (i.e., a large percentage of the total market). A prominent example is a completely nonlocal investor living in New York City or Silicon Valley. The other extreme is typical of investors living in remote areas where there are very few local investment opportunities (and all are very small companies), yet their entire household portfolios are invested locally.

II. Household Locality

A. Household Locality Characteristics

Table II explores the household portfolio locality in still more detail by documenting average portfolio shares across 34,517 households with at least $1,000 in stock holdings at the end of 1991. The first column reports the distribution of household investments by locality that would prevail if investors invested in the most obvious benchmark—the market portfolio. As an alternative, given that the investors are not well diversified (the mean number of stocks in a household's portfolio from our sample is around three), in the second column we consider a different benchmark. The portfolio shares listed for this benchmark represent the hypothetical locality breakdown if, instead of holding its individual stocks, each household invested in the appropriate value-weighted industry portfolio corresponding to each of its individual stocks.[6] The portfolio shares in the second column represent the hypothetical locality breakdown wherein for each stock the measure of locality is the ratio of aggregate market capitalization of all local firms belonging to the same industry and aggregate market capitalization of all firms nationwide belonging to the same industry, with all the individual stock measures then aggregated according to the weights of the respective stocks in the household portfolio. The next five columns present the shares of portfolio value for all households, as well as the subsamples selected according to certain characteristics. Finally, the last column presents the fraction of households that only invest in one of the categories listed in the rows of the table.

The first three columns illustrate how the portfolio shares deviate from the shares that would prevail if every household invested in the market portfolio or the appropriate industry portfolio. For example, the average household invests 32% of its portfolio locally; however, the local share would be only 13% if each household invested in the market or in the appropriate industry portfolio. Thus, relative to either benchmark, individuals' portfolios are subject to a pronounced local bias. The bulk of this local bias can be attributed to investments in local non-S&P 500 stocks. These stocks constitute on average 14% of the household portfolio, but would make up only 3% if the household invested in the market. Households also tend to place disproportionately more weight on local S&P 500 stocks (18% in actual portfolios versus 10% in the market portfolio on average) and substantially less weight on nonlocal S&P 500 stocks (42% in actual portfolio versus 63% in the market portfolio on average). Finally, the average household portfolio share allocated to S&P 500 stocks is 60%, 12% less than the share if the household invested in the market portfolio. Intuitively, it is this 12% under-weighting of the S&P 500 that allows for the 11% over-weighting of nonlocal non-S&P 500 stocks in the average household portfolio. To assess the

[6] Firms are assigned to one of the following 14 industry groups based on their SIC code: mining, oil and gas, construction, food, basic materials, medical/biotechnology, manufacturing, transportation, telecommunications, utilities, retail/wholesale trade, finance, technology, and services. Using either the equally weighted industry portfolio corresponding to each stock or a randomly selected firm in the same industry as the benchmark yields very similar results.

Table II

Composition of Household Portfolios by Locality, 1991

The table reports the average local and nonlocal portfolio shares across 34,517 households with at least $1,000 in stock holdings at the end of 1991. In computing the averages, portfolio shares are weighted by the total household portfolio position in common stocks. Portfolio shares are rounded to one-tenth of a percent. Thus, shares may not add up to 100% due to rounding errors. A stock is defined as local for a given household if the firm is headquartered within 250 miles of the household. Panel A presents the basic breakdown between local and nonlocal stocks, whereas panel B presents a more detailed breakdown that simultaneously considers the stocks' locality and the S&P 500 status. The first column reports the hypothetical shares of household portfolios under the assumption that all households invested in the market portfolio. The second column presents the hypothetical locality breakdown wherein for each stock the measure of locality is the ratio of aggregate market capitalization of all local firms belonging to the same industry and aggregate market capitalization of all firms nationwide belonging to the same industry, with all the individual stock measures then aggregated according to the weights of the respective stocks in the household portfolio. Thus, the portfolio shares listed for this benchmark represent the hypothetical locality breakdown if, instead of holding its individual stocks, each household invested in the appropriate value-weighted industry portfolio corresponding to each of its individual stocks. The third column presents the average portfolio shares for all households in the sample. Tests of statistical significance reported along the first two columns pertain to the test of equality of portfolio shares with portfolios shares in the third column, respectively. The fourth column presents the average portfolio shares of all households except for California households, whereas the fifth column presents the average portfolio shares of California households. Tests of statistical significance reported along the fourth column pertain to the test of equality of portfolio shares with portfolio shares in the fifth column. The sixth and seventh columns present the average portfolio shares of households with portfolio values of at least $10,000 and at least $25,000 (corresponding roughly to the median and the top quartile household portfolio sizes), respectively. The last column presents the fraction of the households in the sample that are invested only in the category highlighted in the respective row of the table.

	If Invested in Market Portfolio	If Invested in Industry Portfolio	All Households	Exclude CA Households	CA Households	Portfolio ≥ $10,000	Portfolio ≥ $25,000	Fraction Invest Only in Category
Panel A: Basic Breakdown								
Local	12.6%*	13.4%*	31.5%	30.6%**	34.0%	31.0%	29.3%	17.0%
Nonlocal	87.4%*	86.6%*	68.5%	69.4%**	66.0%	69.0%	70.7%	48.6%
All stocks	100%	100%	100%	100%	100%	100%	100%	
Panel B: Breakdown with Respect to S&P 500								
Local (S&P 500)	9.5%*	10.0%*	17.6%	17.5%	17.7%	18.6%	18.1%	8.3%
Local (not S&P 500)	3.1%*	3.5%*	13.9%	13.1%**	16.3%	12.5%	11.2%	6.8%
Nonlocal (S&P 500)	62.5%*	60.1%*	42.1%	42.2%	41.8%	45.1%	47.0%	20.2%
Nonlocal (not S&P 500)	24.9%*	26.3%	26.4%	27.2%**	24.2%	23.9%	23.7%	13.3%
All stocks	100%	100%	100%	100%	100%	100%	100%	
Mean number of stocks			3.3	3.4**	3.1	4.3	5.8	
Number of households	34,517	34,517	34,517	25,151	9,366	20,045	9,581	34,517

*Significance of tests of equality of coefficients from the first two columns with the respective coefficient in the third column at the 1% level.

**Significance of tests of equality of coefficients from the Exclude CA Households column with the respective coefficient in the CA households column at the 1% level.

statistical significance of these differences, we carried out significance tests for the first two columns against the third column, respectively. All the differences between the first column, representing the market portfolio, and the third column, representing all households, were statistically significant at the 1% level. Similarly, almost all the differences between the second column, representing the industry-weighted portfolio, and the third column were statistically significant at the 1% level, with the exception being the difference for nonlocal non-S&P 500 stocks.

These tendencies may be related to information asymmetry. Simply put, there are likely to be companies for which there is more information available. Potential determinants of such information asymmetry could include the overall size of the company (measured by market capitalization, number of employees, number of plants, etc.), the number of analysts following the stock, and membership in a major stock index. For example, the potential for sizeable information asymmetry is considerably smaller for S&P 500 companies than it is for non-S&P 500 companies. Thus, it is likely that information asymmetry is most pronounced in the case of nonlocal versus local investment in non-S&P 500 companies. Intuitively, investments in local S&P 500 stocks are unlikely to be driven by the hope to exploit value-relevant information available to local individual investors.

The next four columns of Table II test for robustness of the basic breakdown; they feature virtually identical breakdowns for all households exclusive of California households, for California households, as well as for households with portfolio values of at least $10,000 and at least $25,000 (corresponding roughly to the median and the top quartile household portfolio sizes), respectively. Focusing on households in California, who constitute just over one quarter of the sample, is done to confirm that locality in investments is not driven by these California households. Indeed, while the differences in household portfolio weights of non-California and California households are largely significant (with the exception of local S&P 500 stocks and nonlocal S&P 500 stocks), the magnitudes of the differences are roughly 3% or less. The extent of locality of households with larger portfolios is an empirical issue. It could be higher for smaller portfolios because larger portfolios signal the availability of resources to investigate and uncover value-relevant local information. On the other hand, it could be lower because larger portfolios are likely to consist of more stocks and also are more likely to be combined with investments other than U.S. common stocks, both of which are suggestive of tendencies to achieve better diversification. As it turns out, with the increase in portfolio value, the presence in local non-S&P 500 stocks is slightly reduced and holdings of nonlocal S&P 500 stocks are slightly increased.

Finally, the last column of Table II illustrates that many households are either all local (one-sixth) or all nonlocal (one-half). This is not too surprising, given the small number of stocks in the average household portfolio. About one out of every 15 households holds only non-S&P 500 firms headquartered within 250 miles.

B. Household Locality Determinants

Table III presents the results of fitting cross-sectional regressions of four household portfolio (*HHP*) locality measures, computed at the end of 1991 and summarized in Table I,[7] on several explanatory variables:

$$\begin{aligned} \textit{HHP locality}_i = \alpha &+ \beta_1 \log(\textit{Number of stocks in HHP}_i) \\ &+ \beta_2 \log(\textit{Value of stocks in HHP}_i) \\ &+ \beta_3 \log(\textit{Household income}_i) \\ &+ \beta_4 \textit{Did not report income}_i + \beta_5 \textit{Hold other assets}_i + \varepsilon_i. \end{aligned} \quad (2)$$

Each specification is fitted in both the OLS and the fixed-effects regression frameworks.[8] The intent of the fixed-effect regression, carried out at the zip-code level, is to control for potential biases in the geographic distribution of households and firms in the sample. The first two measures are based on the average distance from the household to its portfolio and thus are negatively related to household portfolio locality. By contrast, the other two measures are based on the fraction of portfolio value within 250 miles from the household and are each positively related to household portfolio locality.

Consistently across all specifications, the variables log(*Number of stocks in HHP*) and *Hold other assets* are statistically significant and are both negatively related to *HHP locality*. This finding is intuitive, as it suggests that a larger extent of diversification translates into a smaller extent of household portfolio locality. On the other hand, log(*Value of stocks in HHP*) is also statistically significant across all specifications, but is positively related to locality. The evidence for log(*Household income*) is somewhat weaker; while it is positively and statistically significantly related with locality in the OLS regressions without fixed effects, it lacks statistical significance in two of the four fixed-effects regressions. Both log(*Value of stocks in HHP*) and log(*Household income*) are potentially indicative of the availability of resources and expertise in information acquisition, which appears to be more efficient in the domain of local stocks, thus resulting in the increase of household portfolio locality.

III. Firm-Level Locality

In this section, we focus on the determinants of local stock ownership. We define $RLO_{j,t}$, the relative local ownership for firm j at the end of year t, as follows:

[7] Results based on year-end measures of locality and explanatory variables for the remaining years in the sample (1992–1995) are qualitatively similar.

[8] To address the potential problems associated with a certain degree of lumpiness in the extent to which households invest locally (the last column of Table II documents that one-sixth of the households in the sample invest only locally, whereas one-half invest only nonlocally), we also employed a Tobit model for all the regressions reported in Table III where the percent of the portfolio invested locally was the dependent variable. The results of these regressions exhibit the same pattern (and levels of significance) as do the reported linear regressions.

Table III

Regression of Locality of Household's Portfolio upon Household and Portfolio Characteristics, 1991

The table presents the results of fitting cross-sectional regressions of four household portfolio (*HHP*) locality measures, computed at the end of 1991, on several explanatory variables according to equation (2) from Section II.B. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. The first two measures are based on the average distance from the household to its portfolio. The other two measures are based on the fraction of portfolio value within 250 miles from the household. Each specification is fitted in both the OLS and the fixed-effects regression frameworks and the corresponding results are presented in the two columns associated with the locality measure. The intent of the fixed-effect regression, carried out at the zip-code level, is to control for potential biases in the geographic distribution of households and firms in the sample. The first two measures are based on the average distance from the household to its portfolio and thus are negatively related to household portfolio locality. By contrast, the other two measures are based on the fraction of portfolio value within 250 miles from the household and are each positively related to household portfolio locality. For those households that did not report income, log(*Household income*) is set to zero and the indicator variable *Did not report income* is set to one. Standard errors (shown in parentheses) allow for heteroskedasticity.

	Average Distance to Portfolio Stocks (Miles)		Average Distance to Portfolio Stocks(Miles)–Distance to Market		Percentage of Portfolio Value within 250 Miles (Local)		Percentage of Portfolio Local–Percentage of Market Local	
log(*Number of stocks in HHP*)	100.0***	117.9***	131.4***	117.9***	−6.7***	−6.7***	−7.5***	−6.7***
	(5.5)	(6.8)	(5.4)	(6.8)	(0.3)	(0.4)	(0.3)	(0.4)
log(*Value of stocks in HHP*)	−16.1***	−22.4***	−32.5***	−22.4***	1.5***	1.4***	1.9***	1.4***
	(3.6)	(4.5)	(3.6)	(4.5)	(0.2)	(0.3)	(0.2)	(0.3)
log(*Household income*)	−10.6**	−6.1	−24.8***	−6.1	2.6***	0.8*	1.6***	0.8*
	(5.1)	(7.0)	(5.1)	(7.0)	(0.3)	(0.4)	(0.3)	(0.4)
Did not report income	5.3	16.6	−58.5***	16.6	7.1***	0.3	3.3**	0.3
	(24.3)	(32.9)	(23.9)	(32.9)	(1.4)	(1.9)	(1.4)	(1.9)
Hold other assets	65.9***	60.6***	71.2***	60.6***	−4.0***	−3.4***	−3.8***	−3.4***
	(7.6)	(9.5)	(7.4)	(9.5)	(0.4)	(0.5)	(0.4)	(0.5)
Constant	1,016***		−42.8		14.1***		2.0	
	(39)		(38.3)		(2.3)		(2.2)	
Zip-code effects	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0167	0.0210	0.0255	0.0210	0.0222	0.0208	0.0249	0.0208
Number of households	34,517	34,517	34,517	34,517	34,517	34,517	34,517	34,517

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

$$RLO_{j,t} = \frac{\$own_{j,t}^{L(j)}}{\$own_{j,t}^{L(j)} + \$own_{j,t}^{NL(j)}} - \frac{\sum_k \$own_{k,t}^{L(k)}}{\sum_k \$own_{k,t}^{L(k)} + \sum_k \$own_{k,t}^{NL(k)}}, \quad (3)$$

where $\$own_{j,t}^{L(j)}$ and $\$own_{j,t}^{NL(j)}$ denote the total value of the position in stock j by the households that had household portfolios of at least \$1,000 and were local and nonlocal to firm j, respectively. As before, the threshold for locality is set to 250 miles. To eliminate the noise stemming from thin ownership in the sample, in the following analyses we include only the firms held by at least five investors in year t. Relative local ownership, $RLO_{j,t}$, is defined as the difference between the fraction of household stock ownership in the firm held by households located within 250 miles of the firm (the first fraction) and the fraction of total nationwide household portfolio wealth held by households located within 250 miles of the firm (the second fraction).

A key result from Section II is that households on average invest disproportionately in local firms. In this section, we consider the dual question of what characterizes a firm's relative local ownership. To that end, we form a panel of all firm-year observations in the sample and regress relative local ownership on several firm characteristics. The panel structure of the data prompts us to account for year effects by adding year dummy variables into the specification and to allow for the correlation of the firm-specific error terms across time.

Table IV presents the regression results for all firms, for S&P 500 firms, and for non-S&P 500 firms. The sample is restricted to stocks owned by at least five households in the sample at the beginning of each year (end of January for 1991) and at the end of November 1996, and consists of seven pooled cross sections: January 1991, December 1992 through 1995, and November 1996.[9] Aside from the regression approach outlined above (denoted as Mean), we also report the results of a median regression estimator (denoted as Median). Regressions fitted for all firms highlight that relative local ownership is related negatively to leverage, positively to the number of employees, and negatively to the S&P 500 membership status, while it is unrelated to the market-to-book ratio, past year's return, and total past year's risk. Indeed, the degree to which local investors might be sensitive to the latter three variables is an empirical issue, but the finding itself is not surprising: It appears that, controlling for other factors, local investments are not overly sensitive to short-term stock performance, nor

[9] In light of the size of the sample, which allows very large t-statistics to develop based on relatively small differences, we checked whether the results of the regressions in Table IV hold for seven separate annual cross-sectional regressions. The results are very robust. For example, for three key significant variables in the pooled regression estimated over all firms—Leverage, Number of employees, and S&P 500 status—we find that all of the annual regression coefficient estimates for these variables have the same sign as the pooled regression estimate and are statistically significant in at least five of the seven annual regressions (with one exception, for which there are four significant years).

Table IV
Relation between Relative Local Ownership (*RLO*) and Firm Characteristics

The table presents the results of fitting cross-sectional regressions of the relative local ownership (*RLO*) measure on several explanatory variables (see Section III for details). *RLO* is defined as the fraction of stock ownership in the firm located within 250 miles of the firm minus the fraction of total nationwide household portfolio value located within 250 miles of the firm (equation (3)). Market value and the Number of employees are each expressed in logarithms. The sample is restricted to stocks owned by at least five households in the sample at the beginning of each year (end of January for 1991) and at the end of November 1996, and consists of seven pooled cross sections from 1991 to November 1996. Regressions are carried out for all firms, S&P 500 firms, and non-S&P 500 firms. For each of the three groups of stocks the table reports results of OLS regressions and median regressions (columns labeled as Mean and Median, respectively). Standard errors (shown in parentheses) allow for heteroskedasticity as well as correlation across observations of the same firm over time. Standard errors for the median regressions are based on 200 bootstrap replications.

	All Firms		S&P 500 Firms		Non-S&P 500 Firms	
	Mean	Median	Mean	Median	Mean	Median
Leverage (debt to assets)	−6.9***	−5.5***	−2.1	−2.0	−7.8***	−7.8***
	(2.0)	(1.2)	(7.4)	(3.9)	(2.0)	(1.5)
Market value	−0.3	0.6***	−1.0	1.1**	−0.2	0.9***
	(0.3)	(0.2)	(0.9)	(0.5)	(0.3)	(0.3)
Market-to-book ratio	−0.0	−0.0	−0.6	−1.2***	−0.0	−0.0
	(0.1)	(0.1)	(0.8)	(0.4)	(0.1)	(0.1)
Number of employees	1.1***	1.1***	−1.0	−1.9***	1.2***	1.5***
	(0.3)	(0.2)	(0.8)	(0.5)	(0.3)	(0.2)
S&P 500	−5.4***	−5.4***				
	(1.2)	(0.6)				
Tech. or biotech.	0.3	2.1***	0.5	2.0**	0.4	2.4***
	(0.8)	(0.6)	(1.6)	(0.9)	(0.9)	(0.6)
Return past 1 year	0.2	−0.1	1.5	−0.1	0.1	−0.2
	(0.3)	(0.3)	(1.0)	(1.3)	(0.3)	(0.3)
Total risk past 1 year	−4.8	−3.2	2.5	26.3**	−3.5	−0.5
	(3.4)	(3.5)	(19.1)	(12.3)	(3.2)	(3.4)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.0088	0.0073	0.0155	0.0075	0.0093	0.0115
Number of observations	16,057	16,057	3,021	3,021	13,036	13,036

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

is there a pronounced difference in local ownership between growth and value stocks. Moreover, technology and biotechnology stocks might be particularly susceptible to high local ownership because the firm performance is highly dependent upon the results of innovation, research, and development of new products, about all of which information may be more readily available locally. The empirical support is mixed, however, as only the median regressions (across all three groups of stocks) suggest a positive relation between firm membership in those two industries and relative local ownership. Finally, relative local ownership is largely unrelated to market value in the mean regressions and is

positively related to it in the median regressions.[10] Firm market value and the number of employees are highly correlated with each other, and if the number of employees is excluded from the mean regression for all firms, the coefficient on firm market value turns positive, though it remains statistically insignificant. From the perspective of information asymmetry, we interpret the number of employees as a noisy proxy of the opportunity to collect value-relevant information locally, or perhaps simply to become aware of the firm's existence (familiarity), which is gained through social interactions with the employees (i.e., by word of mouth).

There is also an alternative explanation for the positive effect of the number of employees on relative local ownership. Other studies have documented the effects of an underlying strong preference for company stock or equities, showing that those who hold a high proportion of equities in pension savings also hold a high fraction of nonpension assets in equities (Bodie and Crane (1997), Weisbenner (2002)), and that employees of firms that require the company match contributions to the 401(k) plan to be invested in company stock tend to invest more of their own 401(k) contributions in company stock (Benartzi (2001), Liang and Weisbenner (2002)). If employees of firms tend to invest part of their nonpension plan wealth in company stock, this could also explain why local ownership is higher for firms with more employees.

However, these issues are unlikely to explain locality in the sample. The majority of the accounts in the database are nonretirement accounts. The accounts that are retirement accounts are either Individual Retirement Accounts (IRAs) or Keough plans; none are retirement plans sponsored by investors' current employers. Thus, a household's investment in its employer's stock through its 401(k) plan is not contained in the data. The only investment in the stock of a company household members currently work for that potentially might be reflected in our data would be investments in non-401(k) accounts. However, if a household member left an employer that offered a 401(k) plan, then the balance of that plan could have been rolled over into an IRA. Thus, while retirement account holdings in our data do not contain stock of the current employer, they may contain stock of a former employer.

Therefore, whereas a key vehicle for own-firm stock ownership, 401(k) plans, is not included in our data, households' potential ownership of a previous employer's stock through an IRA or employees investing in their current firm's stock in nonretirement accounts potentially could cloud the conclusions drawn from our subsequent analyses, as the extent of locality and the returns based on locality might stem primarily from trading on inside information. To address this concern, we conduct appropriate robustness checks in Sections IV.B and IV.F.

[10] In unreported analyses, we compare the distributions of market value of local and nonlocal investments and find that they are very similar. This robustness check alleviates the possible concern that local investments are made into microcap stocks to a larger extent than nonlocal investments, and that many of the results reported in the paper may stem from the extraordinary performance of microcap stocks in the sample period. We thank Josef Lakonishok for pointing out this issue.

We regard S&P 500 status as a fairly accurate measure of both familiarity and the availability of private, value-relevant information that could be more easily accessible to local investors. Simply put, S&P 500 companies are followed by many stock analysts, owned and researched by many institutional investors, held by numerous S&P 500 index funds and other mutual funds, and closely followed by the financial press and other media. Implications of such popularity of S&P 500 stocks are twofold. First, there generally is disproportionately little value-relevant private information regarding S&P 500 stocks to begin with and locality is unlikely to help (living in Seattle likely does not give one a substantial advantage in acquiring information about Microsoft). Second, it does not require locality to be familiar with S&P 500 stocks. Thus, it is not surprising that the loading on the S&P 500 status is significantly negative.

Moreover, the regressions fitted for S&P 500 firms did not yield any statistically significant determinants whose significance would be consistent across the two regressions (mean and median). It thus appears that relative local ownership of S&P 500 firms is not strongly related to firm characteristics. On the other hand, regressions fitted only for non-S&P 500 firms result in the statistical significance, same direction, and comparable magnitude of the impact of leverage and number of employees as do the regressions fitted for all firms. This confirms the intuition that non-S&P 500 firms were the primary drivers of the results from Table IV for all firms.

IV. Information Asymmetry and Locality: Performance and Investment Implications

This section examines the relation between location-driven information asymmetry and household portfolio returns. We seek to quantify the value of locally available information via a simple proposition that individual investors' success in collecting and processing locally available value-relevant information should be reflected in the performance of their household portfolios.

A. *A First Look*

The initial point of inquiry is the relation between household portfolio performance and the allocation of household portfolio weights to local and nonlocal investments. We first regress raw household portfolio returns $R^{HHP}_{t,t+k}$ over k-year horizons ($k = 1, 3$, and 5) on the household portfolio weights of local and nonlocal investment ($\%\, L_t$ and $\%\, NL_t$, respectively) at the end of year t:

$$R^{HHP}_{t,t+k} = \beta_L \%L_t + \beta_{NL} \%NL_t. \tag{4}$$

As before, we control for potential biases in the geographic distribution of households and firms in the sample via the zip-code level fixed-effect regression. The results of fitting regressions based on year $t = 1991$ are presented in the first six columns of Table V (panel A). The difference between the regression

Table V

Relation between Household Portfolio Returns and Locality

The table presents the results of fitting cross-sectional regressions of household portfolio returns on portfolio locality shares (computed at the end of 1991). The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. Household raw returns (calculated based on monthly rebalancing of their portfolio) are calculated over the next year, next 3 years, and next 5 years and are expressed in percentages. These returns are then related to the portfolio shares, computed at the end of 1991, according to the specifications outlined in equations (4) and (5). The results are presented in the first six columns: For each investment horizon, the table reports both the OLS regression and the fixed-effects regression at the zip-code level. The last two columns report results of fitting cross-sectional regressions of average excess monthly household returns α (calculated from a four-factor model according to equation (6) and expressed in basis points) on the portfolio shares computed at the end of 1991 according to the specifications outlined in equation (7). Standard errors (shown in parentheses) allow for heteroskedasticity.

	1 Year Raw Returns (%)		3 Year Raw Returns (%)		5 Year Raw Returns (%)		Average Monthly Excess Returns (bp)	
Panel A: Breakdown by Locality (250 Miles)								
Share of local stocks (L)	13.7***	14.5***	26.4***	29.3***	110.7***	114.3***	17.1***	16.1***
	(0.5)	(0.7)	(1.4)	(2.1)	(3.9)	(7.5)	(1.8)	(2.3)
Share of nonlocal stocks (NL)	10.6***	10.3***	22.3***	20.9***	97.9***	96.1***	10.3***	10.7***
	(0.3)	(0.4)	(0.8)	(1.0)	(2.5)	(3.6)	(1.0)	(1.2)
Locality difference = L − NL	3.2***	4.2***	4.1**	8.4***	12.8***	18.2*	6.8***	5.4*
	(0.7)	(0.9)	(1.7)	(2.8)	(4.9)	(10.5)	(2.2)	(3.1)
Zip-code effects	No	Yes	No	Yes**	No	Yes***	No	Yes
R^2	0.0011	0.0017	0.0006	0.0023	0.0012	0.0023	0.0004	0.0002
Number of observations	27,032	27,032	12,067	12,067	6,462	6,462	26,093	26,093

(*continued*)

Table V—*Continued*

	1 Year Raw Returns (%)		3 Year Raw Returns (%)		5 Year Raw Returns (%)		Average Monthly Excess Returns (bp)	
Panel B: Breakdown by Locality (250 Miles) and S&P 500 Membership								
Share of S&P 500 local stocks (SP, L)	10.8***	11.8***	21.9***	24.5***	100.8***	103.0***	10.8***	9.0***
	(0.5)	(0.8)	(1.4)	(2.6)	(3.9)	(9.3)	(1.9)	(2.7)
Share of S&P 500 nonlocal stocks (SP, NL)	10.1***	9.7***	21.7***	20.3***	99.3***	99.8***	13.8***	14.5***
	(0.3)	(0.4)	(0.8)	(1.4)	(2.4)	(4.9)	(1.1)	(1.6)
S&P 500 locality diff. D(SP) = SP, L − SP, NL	0.8	2.1**	0.2	4.1	1.5	3.1	−3.0	−5.5
	(0.6)	(1.0)	(1.7)	(3.3)	(4.8)	(11.7)	(2.3)	(3.4)
Share of non-S&P 500 local stocks (NSP, L)	17.8***	17.9***	33.5***	36.6***	128.0***	133.2***	25.5***	25.0***
	(1.1)	(1.4)	(2.9)	(4.3)	(8.4)	(15.9)	(3.4)	(4.2)
Share of non-S&P 500 nonlocal stocks (NSP, NL)	11.5***	11.2***	23.4***	22.1***	94.4***	88.4***	4.6**	5.1*
	(0.7)	(0.9)	(2.0)	(2.7)	(7.0)	(10.1)	(2.3)	(2.9)
Non-S&P 500 locality diff. D(NSP) = NSP, L − NSP, NL	6.2***	6.7***	10.1***	14.5***	33.6***	44.8**	20.9***	19.9***
	(1.4)	(1.9)	(3.6)	(5.4)	(11.2)	(19.9)	(4.3)	(5.5)
Difference between the two differences = D(NSP) − D(SP)	5.5***	4.6**	9.9***	10.4	32.1***	41.7*	23.9***	25.4***
	(1.6)	(2.1)	(4.1)	(6.5)	(12.2)	(22.9)	(5.0)	(6.4)
Zip-code effects	No	Yes	No	Yes**	No	Yes***	No	Yes
R^2	0.0031	0.0033	0.0023	0.0042	0.0030	0.0051	0.0017	0.0016
Number of observations	27,032	27,032	12,067	12,067	6,462	6,462	26,093	26,093

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

coefficients on the shares of local and nonlocal investments suggests that local investments outperformed nonlocal ones by 3.2% (4.2% with zip-code level fixed effects) per year on the 1-year horizon; this differential grows over longer investment horizons.[11] Intuitively, investors who invested $x\%$ in local stocks would have outperformed nonlocal investors by $3.2x\%$ ($4.2x\%$) per year on average.

As argued earlier, there likely is disproportionately little value-relevant private information regarding S&P 500 stocks to begin with and locality is unlikely to help. To capture the notion that there should be more information asymmetry between local and nonlocal investors in non-S&P 500 stocks, as opposed to local and nonlocal investors in widely followed and researched S&P 500 stocks, we also estimate the following regression:

$$R^{HHP}_{t,t+k} = \beta_{L,SP}\%L_{SP,t} + \beta_{NL,SP}\%NL_{SP,t} + \beta_{L,NSP}\%L_{NSP,t} + \beta_{NL,NSP}\%NL_{NSP,t}, \tag{5}$$

where $\%L_{SP,t}$ and $\%NL_{SP,t}$ ($\%L_{NSP,t}$ and $\%NL_{NSP,t}$) denote household portfolio weights on local and nonlocal S&P 500 (non-S&P 500) stocks, respectively. Our hypothesis is that the differential in returns between local and nonlocal investments should be more pronounced for non-S&P 500 stocks. The results are presented in the first six columns of Table V (panel B). The difference between the regression coefficients on the shares of S&P 500 local and nonlocal investments suggests that local investments outperformed nonlocal ones by 0.8% (2.1% with zip-code level fixed effects) per year on the 1-year horizon, whereas the differences on longer horizons were not statistically different from zero. By contrast, the difference between the regression coefficients on the shares of non-S&P 500 local and nonlocal investments is substantially higher, that is, 6.2% (6.7% with zip-code level fixed effects) per year on the 1-year horizon; as before, this gap grows with the investment horizon.[12] The seventh row in the panel features the difference between the two differences, which reflects the return to locality in non-S&P 500 companies relative to S&P 500 companies. Its interpretation is the extent to which the premium in investing locally increases with the increase in information asymmetry, in this case 5.5% per year (4.6% with fixed effects) on the 1-year horizon and progressively more for longer investment horizons.[13]

These analyses included all the households that held at least $1,000 in their portfolio at the end of 1991. We test whether larger household portfolios (with portfolio values of at least $10,000 at the end of 1991, which roughly corresponds

[11] Local investments outperformed nonlocal ones by 4.1% (8.4% with fixed effects) over the 3-year horizon and by 12.8% (18.2% with fixed effects) over the 5-year horizon.

[12] Local investments in non-S&P 500 stocks outperformed non-S&P 500 nonlocal ones by 10.1% (14.5% with fixed effects) over the 3-year horizon and by 33.6% (44.8% with fixed effects) over the 5-year horizon.

[13] The return to locality in non-S&P 500 companies relative to S&P 500 companies is 9.9% (10.4% with fixed effects) over the 3-year horizon and 32.1% (41.7% with fixed effects) over the 5-year horizon.

to the median household portfolio size) drive the results from panels A and B because portfolio value may be correlated with the availability of resources and skill in uncovering locality information. The unreported results show that the effect is essentially unchanged for larger portfolios.[14]

To address the concern that the choice of the end of 1991 as the portfolio formation date could have driven the results, we replicate the above analyses for portfolio formation dates in other sample years—at the end of 1992, 1993, 1994, and 1995—for return horizons ranging from 1 year up to 5 years (as permitted by the end of the sample period). Another facet of this robustness check is fitting both the mean estimator (OLS) and the robust median regression estimator. In unreported analyses, we find that the results are very robust along both dimensions: For both mean (OLS) and median estimators, the reported returns to locality for non-S&P 500 stocks are mostly statistically significant, similar in magnitude at a given return horizon, and increasing monotonically with the return horizon.

Another concern we address is the potential relation between the locality of investment and portfolio risk. For example, if local investments were concentrated in small technology stocks, then raw returns for local investments would exceed those for nonlocal investments on average. However, the additional return would reflect a difference in risk (small technology stocks are riskier and thus should generate a higher return on average), as opposed to reflecting the return to the acquisition of local information. Thus, for each household, we compute the average risk-adjusted monthly return by fitting time-series regressions of excess household portfolio returns on four factors (three Fama–French factors and the momentum factor; see Carhart (1997)):

$$R_t^{HHP} - R_{f,t} = \alpha^{HHP} + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \varepsilon_t^{HHP}. \tag{6}$$

We estimate the risk-adjusted household portfolio performance α^{HHP} for all the households in the sample that had at least 24 monthly returns during the 5-year sample period and then regress these estimates (expressed in basis

[14] For example, we find that the difference between the regression coefficients on the shares of non-S&P 500 local and nonlocal investments is 6.2% per year for the 1-year horizon (excess return of 20.9 basis points per month) for the whole sample, and 5.5% (excess return of 27.3 basis points per month) for the sample of households with portfolios of at least $10,000. We further checked whether the relation between locality and returns could be driven by omitted household characteristics that were shown in Table III to be associated with local investing (such as portfolio value and whether the household holds assets other than individual stocks in its portfolio). For example, perhaps higher household wealth could serve as a proxy for the availability of resources and skill to uncover value-relevant local information, and thus its inclusion in the returns regression would render the locality effects insignificant. In unreported analyses, we find the estimated effects of locality to be unchanged with the inclusion of household characteristics such as the size of the household portfolio.

points per month) on the locality breakdown as before:

$$\alpha^{HHP} = \beta_L \%L + \beta_{NL} \%NL \quad \text{and}$$

$$\alpha^{HHP} = \beta_{L,SP} \%L_{SP} + \beta_{NL,SP} \%NL_{SP} + \beta_{L,NSP} \%L_{NSP} + \beta_{NL,NSP} \%NL_{NSP}. \quad (7)$$

The results, presented in the last two columns of Table V, are quite striking: After controlling for risk, across every specification, the best performing investments are local non-S&P 500 stocks and the worst performing investments are nonlocal non-S&P 500 stocks. Two key statistics in panel B—the difference between the regression coefficients on the shares of non-S&P 500 local and nonlocal investments and the difference between differences estimators—suggest that the effect of the information asymmetry of non-S&P 500 local and nonlocal investments in the base specification (without fixed effects) is statistically and economically significant: 20.9 basis points per month for household portfolios that held at least $1,000 at the end of 1991. In terms of magnitude, this coefficient, when annualized, is comparable to those based on raw returns. The difference between the differences, which quantifies the incremental advantage attainable by local investment in non-S&P 500 stocks relative to S&P 500 stocks, is 23.9 basis points per month for household portfolios that held at least $1,000 at the end of 1991. This coefficient is again comparable to those based on raw returns: 23.9 basis points per month translates to approximately 2.9% per year, whereas the corresponding annual return to locality backed out from raw returns was 3.2%. Moreover, risk-adjusted resulted are more robust to the zip-code fixed effects. In sum, adjustment for risk does not affect the main conclusions derived from the earlier analyses based on raw returns.

B. Robustness Checks

Given the time period and the geographic distribution of investors (see Figure 1), a potential concern is that the results may be influenced by certain localities and the success of certain industries. For example, around one quarter of the households in our sample are from California. For the California investors, their local investments could have been in technology companies, which tended to earn superior returns over the sample period. A first look suggests that this fact is unlikely to drive our results. For example, Table II suggests that the portions of portfolios invested locally for California and non-California households were fairly similar. Moreover, the inclusion of the household zip-code fixed effects in our previous regressions controls for potential biases in the geographic distribution of households and firms in the sample. Specifically, in the fixed-effects regressions, the return to locality was estimated on the basis of the differences in portfolio returns within zip codes (rather than on the average differences in returns across the zip codes).

Nonetheless, the importance of this issue merits additional robustness checks. Table VI explores the relation between household portfolio monthly excess returns and locality. For brevity, in panel A we report the difference between local and nonlocal investments, while in panel B we report the difference

Table VI

Robustness Checks on the Relation between Household Portfolio Monthly Excess Returns and Locality

The table presents the robustness checks on the relation between household portfolio monthly excess returns and locality (reported in the last two columns of Table V). The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The table presents key results of the fitting cross-sectional regressions of average excess monthly household returns α (calculated from a four-factor model according to equation (6) and expressed in basis points) on the portfolio shares, computed at the end of 1991, according to the specifications outlined in equation (7). Regressions are fitted for the full sample of households, all households except for California households, California households, and for households that did not have retirement accounts, respectively. For each group of households, the table also documents the regressions that feature portfolio shares across 14 industry groups at the end of 1991. The table reports the differences in estimated coefficients. For the basic locality regression, the top panel, $\beta_L - \beta_{NL}$ is reported. For the regressions that interact locality with S&P 500 membership, displayed in the bottom panel, $\beta_{L,SP} - \beta_{NL,SP}$, $\beta_{L,NSP} - \beta_{NL,NSP}$, and $[(\beta_{L,NSP} - \beta_{NL,NSP}) - (\beta_{L,SP} - \beta_{NL,SP})]$ are reported. Standard errors (shown in parentheses) allow for heteroskedasticity.

	Full Sample		Exclude CA Households		CA Households		Exclude Households With Retirement Accounts	
Panel A: Breakdown by Locality (250 Miles)								
Locality difference = L − NL	6.8***	4.0*	9.4***	8.1***	0.7	−7.4	8.8***	4.3
	(2.2)	(2.2)	(2.5)	(2.5)	(4.4)	(4.6)	(3.1)	(3.1)
Portfolio industry shares included	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0004	0.0217	0.0008	0.0250	0.0000	0.0245	0.0007	0.0267
Number of observations	26,093	26,093	19,063	19,063	7,030	7,030	13,128	13,128
Panel B: Breakdown by Locality (250 Miles) and S&P 500 Membership								
S&P 500 locality diff. D(SP) = SP, L − SP, NL	−3.0	−5.1**	4.6*	2.5	−22.6***	−25.9***	−2.4	−6.4**
	(2.3)	(2.3)	(2.7)	(2.6)	(4.5)	(5.0)	(3.2)	(3.2)
Non-S&P 500 locality diff. D(NSP) = NSP, L − NSP, NL	20.9***	17.3***	16.5***	16.7***	32.1***	16.8**	24.8***	20.0***
	(4.3)	(4.3)	(5.0)	(5.0)	(8.4)	(8.5)	(6.1)	(6.1)
Difference between the two differences = D(NSP) − D(SP)	23.9***	22.4***	11.9**	14.2***	54.7***	42.7***	27.2***	26.4***
	(4.8)	(4.9)	(5.7)	(5.7)	(9.6)	(9.9)	(7.0)	(7.0)
Portfolio industry shares included	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0017	0.0232	0.0012	0.0260	0.0063	0.0286	0.0023	0.0285
Number of observations	26,093	26,093	19,063	19,063	7,030	7,030	13,128	13,128

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

between local and nonlocal S&P 500 investments, the difference between local and nonlocal non-S&P 500 investments, and the difference between the two differences. The first column displays the results from the regression of average excess monthly returns on the share of the portfolio invested locally, thus replicating the results from the second-to-last column in Table V. The second column displays the results when the share of the portfolio allocated across 14 industry groups is also included in the regression. After controlling for the industry composition of the portfolio, the difference between the non-S&P 500 local and nonlocal investments is 17.3 basis points per month, compared to the estimate of 20.9 without this control. The difference between the differences, which quantifies the incremental advantage attainable by local investments in non-S&P 500 stocks relative to S&P 500 stocks, also falls only slightly, from 23.9 to 22.4 basis points, when portfolio industry shares are also included in the regression. This provides compelling evidence that the locality performance results are not driven by particular industries, as the return to locality is still sizeable after controlling for the industry composition of the portfolio.

The next four columns test whether the return to locality is driven by California households. Focusing on the non-California households first, the difference between the non-S&P 500 local and nonlocal investments is 16.5 basis points per month (16.7 when controlling for the industry composition of household portfolios). For California households, the difference between non-S&P 500 local and nonlocal investments is 32.1 basis points per month, substantially larger than that for households elsewhere. This is not particularly surprising, as at the end of 1991, the average California portfolio was tilted 28.6% toward technology stocks, compared to 18.3% for households outside of California. However, once we control for the industry composition of the portfolio, the difference between non-S&P 500 local and nonlocal investments falls to 16.8 basis points per month, in line with our estimates for investors in the rest of the country.

The final two columns display the results when households that had retirement accounts (IRA or Keough plans) are excluded from the sample. This analysis taps into the issue of insider information. Specifically, since households could be holding stock from a previous employer in a retirement account, assuming that the 401(k) plan of the former employer was rolled over into an IRA, the locality of the household's investments could be attributed to holding the former employer's stock. If this were the case, the interpretation of the return to locality could be clouded because it could reflect insider information. Excluding the households with retirement accounts, however, does not diminish the returns to locality; if anything, they are slightly enhanced.

Table VII provides further evidence on whether the return to locality is driven by particular geographic locations of the investors. We break the sample into nine regions, as defined by the U.S. Bureau of the Census (10 regions if we separate California from the Pacific region). For each region, we estimate regressions of both 1-year raw returns and average monthly excess returns and report the difference between the local and nonlocal investments, as well as the difference between the local and nonlocal investments in non-S&P 500 stocks. The results are strikingly consistent across regions. Focusing on the first column, the difference between local and nonlocal investments ranges from a

Table VII
Relation between Household Portfolio Returns and Locality by Region

The table presents the robustness checks on the relation between household portfolio returns and locality (reported in Table V) by region. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. Household raw returns are calculated over the next year and are expressed in percentages. This return is then related to the portfolio shares, computed at the end of 1991, according to the specifications outlined in equations (4) and (5). Aside from regressions involving raw returns, the table presents the results of fitting cross-sectional regressions of average excess monthly household returns α (calculated from a four-factor model according to equation (6) and expressed in basis points) on the portfolio shares, computed at the end of 1991, according to the specifications outlined in equation (7). Households are aggregated by region, as defined by the U.S. Bureau of the Census. Alongside regression results, #HH denotes the number of households in each region. The table reports the differences in estimated coefficients for 1-year raw returns and average monthly excess returns across U.S. regions. For the basic locality regression, reported in the first and fourth columns, the difference $\beta_L - \beta_{NL}$ is reported. For the regressions that interact locality with S&P 500 membership, reported in the second and fifth columns, the difference $\beta_{L,NSP} - \beta_{NL,NSP}$ is reported. Standard errors (not shown) allow for heteroskedasticity.

	1 Year Raw Returns (%)			Average Monthly Excess Returns (bp)		
Region	$\beta_L - \beta_{NL}$	$\beta_{L,NSP} - \beta_{NL,NSP}$	#HH	$\beta_L - \beta_{NL}$	$\beta_{L,NSP} - \beta_{NL,NSP}$	#HH
New England: CT, ME, MA, NH, RI, VT	6.9**	22.0***	1,244	22.3**	15.5	1,232
Middle Atlantic: DE, DC, MD, NJ, NY, PA	2.1	8.1***	4,617	11.9**	10.9	4,456
South Atlantic: FL, GA, NC, SC, VA, WV	9.8***	8.6*	2,949	7.8	−1.6	2,804
East north central: IL, IN, OH, MI, WI	0.2	2.3	3,078	−21.0***	−24.6**	2,911
East south central: AL, KY, MS, TN	7.6**	17.1***	661	−1.0	52.7**	672
West north central: IA, KS, MN, MO, NE, ND, SD	−2.4	0.2	1,342	15.9	7.4	1,285
West south central: AR, LA, OK, TX	10.9***	14.9***	2,488	7.6	7.5	2,411
Mountain: AZ, CO, ID, MT, NV, NM, UT, WY	7.4*	4.6	1,738	10.4	26.7	1,743
Pacific: CA, OR, WA	1.9	4.1	8,915	10.4***	43.3***	8,579
Pacific—exclude CA	−7.8***	−9.0	1,634	49.3***	77.7***	1,549
CA	4.5***	8.4***	7,281	0.7	32.1***	7,030
All regions	3.2***	6.2***	27,032	6.8***	20.9***	26,093

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

low of −2.4 in the West north central region, to 10.9 in the West south central region. Eight of the nine regions have positive returns to investing locally, with statistically significant results in five regions. For all nine regions, local non-S&P 500 investments outperform nonlocal non-S&P 500 investments, with the difference being statistically significant in five. The differential in 1-year returns exceeds 14% in the East south central, the West south central, and New England. A similar pattern emerges when we focus on excess returns: Local investments outperform nonlocal investments in the majority of the regions, and often significantly so.

Throughout the sample period, households enter and leave the sample. Of the 34,517 households with at least $1,000 in stock holdings at the end of 1991, the database reported an equity position for around one-fourth of those households in November 1996. Other households either appear in the sample after a certain number of months or disappear from the sample by liquidating their positions and closing the account. Thus, in a given period we observe the performance of a household portfolio if it "survives," that is, if the household's accounts do not get liquidated. A potential concern at this point might be survivorship bias (see, e.g., Brown et al. (1992)). Indeed, in the domain of performance measurement of professional money managers, survivorship bias is a concern because money managers who disappear from the sample have often been underperformers and their absence from the sample biases upward performance measures of the surviving managers. In our sample, we find no evidence that individual investors systematically liquidated their positions because of their underperformance.[15]

The final concern is that the portfolio breakdown with respect to locality, that is, the weights $\%L_{SP}$, $\%NL_{SP}$, $\%L_{NSP}$, and $\%NL_{NSP}$, are computed at the beginning of the returns measurement period. Although in unreported analyses we find that household locality is a highly persistent phenomenon, household portfolio locality may change over time, and relying on the portfolio locality breakdown at the beginning of the return measurement period introduces noise that may worsen as the returns period increases. While this

[15] To test whether the liquidation of accounts was related to household portfolio performance, we calculate the portfolio returns of both the households that remained in the sample until the end (November 1996) and those who liquidated their account before the end of the sample period. There were 27,032 households that appeared in the regression relating the locality of their portfolio at the end of 1991 to portfolio returns in 1992. Of these households, the 6,462 that maintained stock ownership in their accounts continuously through November 1996 had an average portfolio return of 10.7% in 1992. The remaining 20,570 households, which liquidated their accounts some time before the end of the sample period, had an average portfolio return of 11.9% in 1992. Furthermore, there were 26,093 households for which we calculate average monthly excess (risk-adjusted) returns and relate them to household locality. The average excess monthly return for the households that maintained stock ownership in their accounts continuously through November 1996 was 7.9 basis points, compared to 16.4 basis points for the households that liquidated their account some time before the end of the sample. Finally, in unreported analyses, we regress the probability of a household liquidating its account during the year to the household's prior 1-year portfolio return separately for each year. In no year was there a significantly negative relation between the prior portfolio return and the probability of liquidating the account. In sum, we do not uncover evidence that poor household returns led to household portfolio liquidation.

consideration may have some bearing on the precision of the estimates pertaining to longer horizons, the 1-year measures of information advantages of locality are unlikely to be substantially affected by the potentially stale portfolio locality breakdown. Moreover, it is not clear how to avoid stale information in this context, as regressing returns on some type of average portfolio locality breakdown computed *during* the return measurement period would raise concerns about endogeneity. We alleviate this concern in the next section by forming portfolios that include households with certain profiles of local investment and rebalance such portfolios annually.

C. *Pursuing Locality: The Portfolio Perspective*

The previous section suggests that household portfolio performance is positively related to the extent and nature of locality in investors' portfolios and that the bulk of the performance differential stems from the investments in less widely known (non-S&P 500) stocks. In this and the next several sections, we seek to quantify the economic value of local investment. We begin with a simple question: What are the benefits of pursuing locality by simply following the household portfolios aggregated according to varying degrees of their locality of investment?

At the beginning of each year (end of January for 1991), households are ranked based on the locality of their portfolio relative to the market portfolio (the fraction of household portfolio within 250 miles minus the fraction of market portfolio within 250 miles) and then assigned to locality quartiles. Average household relative locality cutoffs across all years are −6.3% for the first quartile, −0.5% for the second quartile, and 46.6% for the third quartile. The return for each quartile represents the weighted return for all the households in that quartile (the weight is the value of each household's portfolio). Thus, these four locality portfolios are rebalanced at the beginning of each year based on households' locality relative to the market portfolio. The sample period is from February, 1991 to December, 1996.

Table VIII summarizes the returns to the four portfolios ranked by the household portfolio locality. Panel A presents the average monthly raw returns and panel B presents the risk-adjusted returns, computed as excess monthly returns from the same four-factor model we used in the previous section (see equation (6)). For both raw returns and risk-adjusted returns, the bottom three household locality quartile portfolios have similar performances and only the top quartile portfolio stands out. The performance differential relative to the top quartile, highlighted in the second column of each panel, is 19–24 basis points per month for raw returns and 14–16 basis points for risk-adjusted returns. This translates into the top household locality quartile outperforming the remaining three portfolios by 34–41% (26–29%) during the sample period according to the raw (risk-adjusted) portfolio returns presented in the table. The evidence across both panels not only stresses the performance superiority of the portfolio of most local households, but also shows that this performance differential is largely not attributable to differences in risk exposures across portfolios.

Table VIII

Performance of Household Portfolios Aggregated by Locality

The table documents the performance of household portfolios aggregated into four portfolios according to their locality. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. At the beginning of each year (end of January 1991), households are ranked based on the locality of their portfolio relative to the market portfolio (fraction of household portfolio within 250 miles minus the fraction of market portfolio within 250 miles) and then assigned to locality quartiles. The sample period is from February 1991 to December 1996. The average household relative locality cutoffs across all years are −6.3% for the first quartile, −0.5% for the second quartile, and 46.6% for the third quartile. The return for each quartile represents the weighted return for all the households in that quartile (the weight is the value of each household's portfolio). Raw returns are presented in panel A, whereas excess returns are presented in panel B. Excess monthly returns are calculated from a four-factor model, which accounts for the market, size, book-to-market, and momentum effects (Carhart (1997)). These four locality portfolios are rebalanced at the beginning of each year based on households' locality relative to the market portfolio. In each panel, aside from reporting the performance of the four locality portfolios, for the bottom three locality-quartile portfolios, the table also reports their performance differential relative to the most local quartile portfolio. Standard errors reported are Newey-West standard errors with autocorrelation up to three lags.

	Panel A: Raw Returns (%)		Panel B: Excess Returns (%)	
Household by Locality Quartile	Average Monthly Return	Difference Relative to Top Quartile	Average Excess Monthly Return	Difference Relative to Top Quartile
Quartile 1 (least local)	1.393***	−0.189	−0.003	−0.157
	(0.375)	(0.120)	(0.108)	(0.103)
Quartile 2	1.344***	−0.238**	0.016	−0.138
	(0.333)	(0.122)	(0.084)	(0.103)
Quartile 3	1.361***	−0.221*	0.015	−0.140
	(0.331)	(0.125)	(0.080)	(0.103)
Quartile 4 (most local)	1.582***		0.154	
	(0.425)		(0.161)	

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

D. Pursuing Locality: Dissecting Portfolios by Locality

The analyses from the previous section rely on a simple portfolio strategy that gauges the broad impact of locality on household portfolio performance. In this section, we proceed a step further by first dissecting each household portfolio according to two dimensions, locality of the investments and their S&P 500 status (our proxy for information asymmetry), and then forming six aggregate position-weighted portfolios across all households.[16]

The results are presented in Table IX. The first section of the table (left three columns) is based on the portfolios characterized only by the locality of the securities relative to the households that held them, whereas the next two sections recognize not only locality, but also the S&P 500 status. Each section

[16] Our approach is similar in spirit to that followed by Coval and Moskowitz (2001).

Table IX
Regression of Monthly Portfolio Returns from February 1991 to December 1996, Local versus Nonlocal Investments

The table documents the performance of aggregate household portfolios constructed by aggregating all the households' local holdings and nonlocal holdings. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The sample period is from February 1991 to December 1996. The average raw return per month and the excess return per month (Alpha) from various portfolios are expressed in percentage points. Excess monthly returns are calculated from a four-factor model, which accounts for the market, size, book-to-market, and momentum effects (Carhart (1997)). The loadings on each of the four factors are presented for nine portfolios. Portfolio 1 (All Stocks, Local) is the return on local investments weighted across all household portfolios. Portfolio 2 (All Stocks, Nonlocal) is the return on nonlocal investments weighted across all household portfolios. Portfolio 3 (All Stocks, Difference) is the return on the zero-cost portfolio that is long local stocks (Portfolio 1) and is short nonlocal stocks (Portfolio 2), again weighted across all household portfolios. The middle three portfolios are defined analogously, but just focus on S&P 500 stocks held by households. The right three portfolios are defined analogously, but just focus on non-S&P 500 stocks held by households. Standard errors reported are Newey-West standard errors with autocorrelation up to three lags.

	All Stocks			S&P 500 Stocks			Non-S&P 500 Stocks		
	Local	Nonlocal	Difference	Local	Nonlocal	Difference	Local	Nonlocal	Difference
Market	1.14***	1.07***	0.07*	1.06***	1.03***	0.03	1.26***	1.15***	0.11*
	(0.06)	(0.03)	(0.04)	(0.06)	(0.03)	(0.03)	(0.08)	(0.06)	(0.06)
Small minus big	0.24***	0.15***	0.09**	−0.14**	−0.17***	0.03	0.84***	0.84***	−0.00
	(0.09)	(0.05)	(0.04)	(0.07)	(0.05)	(0.03)	(0.11)	(0.08)	(0.06)
High minus low	−0.09	−0.10**	0.01	−0.08	−0.09*	0.01	−0.11	−0.12***	0.01
	(0.07)	(0.05)	(0.03)	(0.08)	(0.05)	(0.04)	(0.07)	(0.05)	(0.04)
Up minus down	−0.08	−0.07**	−0.01	−0.11*	−0.12***	0.01	−0.05	0.03	−0.07
	(0.06)	(0.04)	(0.03)	(0.06)	(0.04)	(0.03)	(0.09)	(0.05)	(0.05)
p-value of regression	0.000***	0.000***	0.006***	0.000***	0.000***	0.511	0.000***	0.000***	0.305
Alpha	0.10	0.02	0.09	0.22*	0.16*	0.06	−0.06	−0.26	0.20**
	(0.13)	(0.09)	(0.07)	(0.15)	(0.09)	(0.07)	(0.18)	(0.17)	(0.10)
Raw return	1.53***	1.36***	0.17**	1.48***	1.37***	0.11	1.61***	1.36***	0.25**
	(0.39)	(0.34)	(0.08)	(0.35)	(0.31)	(0.07)	(0.53)	(0.49)	(0.11)

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

features the performance of portfolios of local and nonlocal stocks, drawn from the appropriate universe of stocks (all, S&P 500, and non-S&P 500), as well as the key portfolio of interest—the zero-cost portfolio formed by taking the long position in the portfolio of local stocks and the short position in the portfolio of nonlocal stocks.

The results support the hypothesis that locally available information is value-relevant. Specifically, the zero-cost portfolio based on all stocks has statistically and economically significant returns: Monthly raw returns are 17 basis points (9 basis points after adjusting for risk). The next two sections of Table IX further demonstrate that this effect stems primarily from the stocks that exhibit more information asymmetry. Indeed, the corresponding raw returns and risk-adjusted returns in the next two sections are very different: They are substantially smaller for S&P 500 stocks, which most likely do not feature significant information asymmetries by investors' location, whereas they are considerably larger (25 and 20 basis points per month, respectively) precisely where information asymmetries come into play—in the domain of non-S&P 500 stocks.

These results are striking, as the only portfolio formation criteria are related to geography and a basic stock index membership characteristic. Even these simple breakdowns yield abnormal performance of 2.4–3.0% per year, and adjustment for risk characteristics of the portfolios makes little difference.

E. Pursuing Locality: Combining the Two Perspectives

The results from the previous section, while compelling in their own right, do not take into account the fact that some households had very few local investments, whereas others had at least one-half of their portfolio invested in local stocks.[17] In Section IV.C, we found that households that invest locally tend to outperform the rest, although we did not determine the source of their relative success. In this section, we advance our analyses a step further by bringing the two inquiries together.

The question at hand is simple: Were the investors who invested more locally better able to uncover and exploit locally available value-relevant information than the investors who invested less locally? To address it, we split each of the six portfolios formed in the previous section into two subportfolios. The first (second) portfolio consists of all the portfolio positions held by households ranked below (above) the median of the household locality distribution.[18]

At the top of Table X, we reproduce the key statistics highlighted near the bottom of Table IX, namely the raw and risk-adjusted average monthly returns for the zero-cost portfolios formed by taking the long (short) position in the respective local (nonlocal) portfolio. The next two rows summarize the analogous returns for the portfolios based on the positions held by below-median locality and above-median locality households. The results show that the ability to

[17] As documented in Section IV.C, the top quartile of households with respect to locality has at least 46.6% of their portfolio value invested locally.

[18] The average of the household relative locality medians across all years is −0.5%.

Table X

Average Monthly Returns of Zero-Cost Portfolios (Local Minus Nonlocal) from February 1991 to December 1996, Formed with Respect to Household Portfolio Locality

The table documents the performance of aggregate household portfolios constructed by aggregating all the households' local holdings and nonlocal holdings while controlling for household locality. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The sample period is from February 1991 to December 1996. The average raw return per month and the excess return per month (Alpha) from various portfolios are expressed in percentage points. Excess monthly returns are calculated from a four-factor model, which accounts for the market, size, book-to-market, and momentum effects (Carhart (1997)). The left panel (All Stocks) focuses on all stocks held by households, and breaks those holdings into local and nonlocal investments. The return on the zero-cost portfolio that is long local stocks and is short nonlocal stocks weighted across all household portfolios is displayed in the top row. This calculation is then repeated restricting the sample to households whose portfolio locality share is below the median (middle row) and above the median (bottom row). At the beginning of each year (end of January 1991), households are ranked based on the locality of their portfolio relative to the market portfolio (fraction of household portfolio within 250 miles minus fraction of market portfolio within 250 miles). The average household relative locality median cutoff across all years is −0.5% for the median. The three rows of the middle panel (S&P 500 Stocks) are defined analogously, but just focus on S&P 500 stocks held by households. The three rows of the right panel (Non-S&P 500 Stocks) are defined analogously, but just focus on non-S&P 500 stocks held by households. Standard errors reported are Newey-West standard errors with autocorrelation up to three lags.

	All Stocks		S&P 500 Stocks		Non-S&P 500 Stocks	
	Raw Return	Alpha	Raw Return	Alpha	Raw Return	Alpha
All households	0.17**	0.09	0.11	0.06	0.25**	0.20**
	(0.08)	(0.07)	(0.07)	(0.07)	(0.11)	(0.10)
Households by locality						
Household locality	0.00	−0.06	0.06	0.09	0.01	−0.06
below median	(0.09)	(0.09)	(0.10)	(0.11)	(0.18)	(0.18)
Household locality	0.22**	0.12	0.12	0.05	0.39***	0.28***
above median	(0.09)	(0.08)	(0.08)	(0.08)	(0.11)	(0.10)

*** and ** indicate significance at the 1% and 5% levels, respectively.

detect and exploit locally available value-relevant information is concentrated in the households whose portfolios tended to be more local. In each of the six columns, the second row, based on below-median locality households, features zero-cost portfolio average returns indistinguishable from zero.

In almost all cases, zero-cost portfolio average returns in the third row, based on above-median locality households, are substantially larger in magnitude than the portfolio returns based on below-median locality households. In particular, the last two columns that pertain to the zero-cost portfolios most affected by the existence of locally available value-relevant information—those formed on the basis of non-S&P 500 stocks—showcase the highest average monthly returns: 0.39% for raw returns and 0.28% for risk-adjusted returns.

These figures can be interpreted from the perspective of seeking to exploit the information contained in the investors' positions: Paying attention not only to the stock characteristics, but also to the extent of household locality, can enhance average zero-cost portfolio raw (risk-adjusted) returns from 0.25% (0.20%) to 0.39% (0.28%). Moreover, this improvement, though moderate, is attainable without any incremental cost except for screening by household locality.

F. Further Robustness Checks

Tables XI and XII provide additional evidence of the robustness of our results. Recall that a key vehicle for own-firm stock ownership, 401(k) plans, is not included in our data. A household's potential ownership of a previous employer's stock through an IRA or an employee investing in their current employer's stock in nonretirement accounts, however, could cloud our conclusions, as the results could be driven by the (former) employees trading on inside information. To address this concern, in Section IV.B we excluded households with retirement accounts and found that such exclusion did not affect our results. In Table XI, we explore the performance of a household's largest local investment relative to its other, if any, local investments. The motivation for this analysis is intuitive: If the household were holding its own-employer's stock, it would likely be its largest local holding. Thus, if households purchase their own employer's stock based on inside information, the return on the portfolio of these biggest local investments, aggregated across households, should outperform the aggregated portfolio of all other local holdings across households.[19]

Across all stocks, across S&P 500 stocks only, as well as across non-S&P 500 stocks only, we cannot reject the hypothesis that there is no difference in performance in raw returns between a household's largest local investment and its other local investments. The same pattern emerges for risk-adjusted returns, with the largest local investment actually slightly underperforming the other local investments. Also, in unreported analyses, we find that, consistent with Table IX, the difference in returns in the zero-cost portfolio that takes a long position in the local investments exclusive of the biggest local investment and a short position in the nonlocal investments offers a significantly positive return, a finding that is even stronger when the portfolios contain only non-S&P 500 investments. In sum, these results provide compelling evidence that the performance of households' local investments is not driven by insider trading, as there is no difference in the performance of the largest local holding relative to other local investments.

Table XII provides further evidence on the return to locality for California households relative to investors elsewhere, augmenting the results in Tables VI

[19] On average, households hold three stocks, one of which is the stock of a local company. Thus, it is not surprising that the household's biggest local investment, which may be its only local stock, comprises on average 27.7% of a household's portfolio at the end of 1991, compared to 3.8% for the other local investments.

Table XI

Regression of Monthly Local Stock Portfolio Returns from February 1991 to December 1996, Biggest Local versus Other Local Investments

The table documents the performance of aggregate local household portfolios. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The sample period is from February 1991 to December 1996. The average raw return per month and the excess return per month (Alpha) from various portfolios are expressed in percentage points. Excess monthly returns are calculated from a four-factor model, which accounts for the market, size, book-to-market, and momentum effects (Carhart (1997)). The loadings on each of the four factors are presented for nine portfolios. Portfolio 1 (All Local Stocks, Biggest Local Stock) is the return on the biggest local investment weighted across all household portfolios. Portfolio 2 (All Local Stocks, Other Local Stocks) is the return on local investments, other than the household's largest, weighted across all household portfolios. Portfolio 3 (All Local Stocks, Difference) is the return on the zero-cost portfolio that is long the biggest local stock (Portfolio 1) and is short the other local stocks (Portfolio 2), again weighted across all household portfolios. The middle three portfolios are defined analogously, but just focus on S&P 500 stocks held by households. The right three portfolios are defined analogously, but just focus on non-S&P 500 stocks held by households. Biggest S&P 500 Local Stock represents the household's biggest local investment if it happens to be a member of the S&P 500, and Biggest Non-S&P 500 Local Stock represents the household's biggest local investment if it happens to not be a member of the S&P 500. Standard errors reported are Newey-West standard errors with autocorrelation up to three lags.

	All Local Stocks			S&P 500 Local Stocks			Non-S&P 500 Local Stocks		
	Biggest	Other	Difference	Biggest	Other	Difference	Biggest	Other	Difference
Market	1.14***	1.15***	−0.01	1.05***	1.10***	−0.05***	1.28***	1.22***	0.06
	(0.06)	(0.05)	(0.02)	(0.06)	(0.05)	(0.02)	(0.09)	(0.07)	(0.04)
Small minus big	0.25***	0.23***	0.01	−0.13*	−0.17**	0.03	0.83***	0.87***	−0.04
	(0.09)	(0.09)	(0.02)	(0.07)	(0.08)	(0.03)	(0.12)	(0.12)	(0.04)
High minus low	−0.11	−0.04	−0.07***	−0.09	−0.05	−0.04	−0.15*	−0.02	−0.14***
	(0.07)	(0.07)	(0.02)	(0.08)	(0.08)	(0.03)	(0.08)	(0.07)	(0.04)
Up minus down	−0.06	−0.14***	0.08***	−0.09	−0.18***	0.09***	−0.03	−0.10	0.07*
	(0.06)	(0.06)	(0.02)	(0.06)	(0.06)	(0.02)	(0.09)	(0.08)	(0.04)
p-value of regression	0.000***	0.000***	0.000**	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Alpha	0.07	0.19*	−0.11*	0.19	0.30***	−0.11	−0.08	0.02	−0.10
	(0.14)	(0.11)	(0.07)	(0.14)	(0.11)	(0.07)	(0.20)	(0.17)	(0.10)
Raw return	1.51***	1.59***	−0.08	1.46***	1.55***	−0.09	1.60***	1.65***	−0.05
	(0.40)	(0.38)	(0.07)	(0.35)	(0.34)	(0.07)	(0.54)	(0.51)	(0.10)

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table XII

Regression of Monthly Portfolio Returns from February 1991 to December 1996, Robustness to California

The table presents a robustness check of the performance of aggregate household portfolios (presented in Table IX) with respect to the California households. The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The sample period is from February 1991 to December 1996. The average raw return per month and the excess return per month (Alpha) from various portfolios are expressed in percentage points. Excess monthly returns are calculated from a four-factor model, which accounts for the market, size, book-to-market, and momentum effects (Carhart (1997)). The loadings on each of the four factors are presented for six portfolios. Portfolio 1 (Non-S&P 500 Stocks, Local) is the return on local, non-S&P 500 local investments weighted across all non-California household portfolios. Portfolio 2 (Non-S&P 500, Non-local) is the return on nonlocal, non-S&P 500 investments weighted across all non-California household portfolios. Portfolio 3 (Non-S&P 500 Stocks, Difference) is the return on the zero-cost portfolio that is long non-S&P 500 local stocks (portfolio 1) and is short nonlocal stocks (portfolio 2), again weighted across all non-California household portfolios. The right three portfolios are defined analogously, but just focus on non-S&P 500 stocks held by the California households. Standard errors reported are Newey-West standard errors with autocorrelation up to three lags.

	Exclude CA Households, Non-S&P 500 Stocks			CA Households, Non-S&P 500 Stocks		
	Local	Nonlocal	Difference	Local	Nonlocal	Difference
Market	1.11***	1.16***	−0.05	1.54***	1.12***	0.42***
	(0.05)	(0.06)	(0.05)	(0.18)	(0.07)	(0.15)
Small minus big	0.79***	0.82***	−0.04	0.98***	0.87***	0.11
	(0.06)	(0.08)	(0.06)	(0.24)	(0.08)	(0.19)
High minus low	−0.04	−0.11**	0.07	−0.23	−0.17**	−0.06
	(0.07)	(0.05)	(0.07)	(0.19)	(0.05)	(0.18)
Up minus down	−0.02	0.03	−0.05	−0.08	0.02	−0.09
	(0.07)	(0.05)	(0.06)	(0.19)	(0.05)	(0.15)
p-value of regression	0.000***	0.000***	0.317	0.000***	0.000***	0.000***
Alpha	0.10	−0.25	0.35*	−0.30	−0.29	−0.01
	(0.18)	(0.17)	(0.20)	(0.49)	(0.19)	(0.43)
Raw return	1.67***	1.39***	0.29*	1.59**	1.27**	0.32
	(0.47)	(0.49)	(0.16)	(0.75)	(0.51)	(0.40)

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

and VII. Focusing on non-S&P 500 investments, we seek to establish whether the excess return of 20 basis points per month to the non-S&P 500 zero-cost portfolio (long local and short nonlocal), shown in Table IX, is driven by California investors. Both risk-adjusted performance and raw returns to the zero-cost strategies suggest that California investors did not outperform non-California ones. If anything, risk-adjusted performance seems to be larger for non-California investors.

As a final robustness check, we separate investments in technology and medical/biotechnology stocks, two particularly good performers during our sample period, from all other investments, and test whether the superior performance of local investments held for both groups of investments. At the end of 1991, technology/medical stocks comprised on average 35.4% of the household portfolio. We find that, focusing on non-S&P 500 investments, the return to the zero-cost portfolio (aggregated across all households) that takes a long position in local tech/medical stocks and a short position in nonlocal tech/medical stocks was 32 basis points per month (p-value = 0.146), while the return on the analogous zero-cost portfolio that excludes tech/medical stocks was 23 basis points per month (p-value = 0.039). The difference is not statistically different from zero (p-value = 0.741). This analysis complements the finding in Table VI that the return to locality persists after controlling for the industry composition of the portfolio.

G. Pursuing Locality: Follow the Locals

As an indirect test of the value of locally available value-relevant information, we examine more closely the performance of the nonlocal non-S&P 500 part of a household's portfolio. Non-S&P 500 stocks that are distant to a household are the stocks for which the household likely is at the greatest information disadvantage relative to local investors. Suppose one considers investing in the stock of a remote, nonlocal firm. Further, suppose that there is locally available value-relevant information about the firm. If it turns out that the locals invest heavily in this stock, chances are that the stock will perform well. On the other hand, if locals (who likely have an information advantage on the prospects of the firm, particularly for non-S&P 500 companies) shy away from the stock, chances are that the performance will be lackluster. Thus, if equipped with the knowledge about a remote stock's relative local ownership, a nonlocal investor might be able to reap the benefits of the local investors' due diligence. To capture the relative local ownership effect, we employ the regression analyses similar to those used in Section IV.A. Specifically, Table XIII presents the results of fitting the following regressions:

$$
\begin{aligned}
R^{HHP}_{t,t+k} &= \beta_1 \%L_{SP,t} + \beta_2 \%NL_{SP,t} + \beta_3 \%L_{NSP,t} \\
&\quad + \beta_4 \%NL^{High\,RLO}_{NSP,t} + \beta_5 \%NL^{Low\,RLO}_{NSP,t} \quad \text{and} \\
\alpha^{HHP} &= \beta_1 \%L_{SP} + \beta_2 \%NL_{SP} + \beta_3 \%L_{NSP} \\
&\quad + \beta_4 \%NL^{High\,RLO}_{NSP} + \beta_5 \%NL^{Low\,RLO}_{NSP},
\end{aligned}
\tag{8}
$$

Table XIII

Relation between Household Portfolio Raw Returns and Locality Revisited (Relation between Returns and Nonlocal Share Interacted with *RLO*)

The table presents the results of fitting cross-sectional regressions of household portfolio returns on portfolio locality shares (computed at the end of 1991) with a particular emphasis on relative local ownership (*RLO*), as defined in equation (3). The sample consists of 34,517 households with at least $1,000 in stock holdings at the end of 1991. A stock is local for a given household if the firm is headquartered within 250 miles of the household. The median *RLO* cutoff is 9.8%. On average, non-S&P 500 nonlocal stocks constitute 26.4% of the value of a household's portfolio (13.4% has high *RLO*, and 13.0% has low *RLO*). Household raw returns (calculated based on monthly rebalancing of their portfolios) are calculated over the next year, next 3 years, and next 5 years and are expressed in percentages and average excess monthly household returns α (calculated from a four-factor model according to equation (6)) are expressed in basis points. These returns are then related to the portfolio shares, computed at the end of 1991, according to the specifications outlined in equation (8). Standard errors (shown in parentheses) allow for heteroskedasticity.

	Raw Returns			Average Excess
	1 Year	3 Years	5 Years	Monthly Returns
Share of S&P 500 local stocks	10.8***	21.9***	100.8***	0.11***
	(0.5)	(1.4)	(3.9)	(0.02)
Share of S&P 500 nonlocal stocks	10.0***	21.7***	99.2***	0.14***
	(0.3)	(0.8)	(2.4)	(0.01)
Share of non-S&P 500 local stocks	17.8***	33.5***	128.1***	0.26***
	(1.1)	(2.9)	(8.4)	(0.03)
Share of non-S&P 500 nonlocal stocks with high relative local ownership (high *RLO*)	15.7***	28.3***	115.7***	0.15***
	(1.0)	(2.8)	(10.4)	(0.03)
Share of non-S&P 500 nonlocal stocks with low relative local ownership (low *RLO*)	7.3***	18.8***	73.7***	−0.07*
	(1.2)	(3.0)	(11.0)	(0.04)
Non-S&P 500 local − non-S&P 500 nonlocal (high *RLO*)	2.0	5.2	12.4	0.11**
	(1.6)	(4.1)	(13.8)	(0.05)
Non-S&P 500 nonlocal (high *RLO*) − non-S&P 500 nonlocal (low *RLO*)	8.4***	9.5**	42.0***	0.21***
	(1.7)	(4.3)	(16.3)	(0.05)
p-value: All non-S&P 500 coefficients the same	0.000***	0.003***	0.001***	0.000***
R^2	0.0049	0.0030	0.0054	0.0028
Number of observations	27,032	12,067	6,462	26,093

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

where, as before, $\%L_{SP,t}$ and $\%NL_{SP,t}$ ($\%L_{NSP,t}$ and $\%NL_{NSP,t}$) denote household portfolio weights on local and nonlocal S&P 500 (non-S&P 500) stocks, respectively. In this specification, the latter is further split according to relative local ownership into $\%NL_{NSP,t}^{HighRLO}$ and $\%NL_{NSP,t}^{LowRLO}$.[20]

Regardless of the returns horizon or risk-adjustment, the hypothesis that all coefficients on non-S&P 500 stocks are the same is easily rejected. At the same time, the difference between the raw returns on local non-S&P 500 stocks and nonlocal non-S&P 500 investments with high relative local ownership is not statistically different from zero over any return horizon. The risk-adjusted return of 11 basis points per month translates into approximately 1.3% per year. Thus, merely favoring nonlocal non-S&P 500 stocks with high relative local ownership goes a long way toward reducing and almost eliminating the information asymmetry between the local and nonlocal non-S&P 500 stocks. The second highlighted row presents the difference between investing in nonlocal non-S&P stocks with high and low relative local ownership, respectively. A risk-adjusted monthly differential of 21 basis points translates into the annual differential of around 2.5%, which, though considerably smaller than the raw return differential of 8.4%, is still substantial. Another interpretation of the 21 basis point differential is that it is comparable to the risk-adjusted return resulting from the zero-cost portfolio strategies on non-S&P 500 stocks (20 basis points without controlling for the participating household portfolio locality and 28 basis points for the zero-cost portfolio strategies involving only the households with above-median locality). In sum, mimicking what the locals do is about as good as being local (if you cannot be local, join the locals), and betting that you may have found a winning investment that the local investors have overlooked more often than not is a losing proposition.

H. Something Closer

Our choice of 250 miles as the cutoff between local and nonlocal investments was governed by an assumption that the distance of 250 miles is a plausible upper bound on the span of local information. As discussed earlier, this cutoff is admittedly arbitrary, and it could be argued that it is on the conservative side. As a robustness check, in unreported analyses we replicated the basic analyses from Table V with the locality cutoff set at 100 km (around 62 miles).

Changing the locality threshold has several implications. First, the breakdown into local and nonlocal portfolio shares changes from 32% versus 68% to 21% versus 79%.[21] Second, the results based on the 100 km-cutoff are very consistent, and are even stronger by 20–50%, which suggests that our choice of 250 miles as the cutoff indeed was conservative.

[20] On average, non-S&P 500 nonlocal stocks constitute 26.4% of the value of a household's portfolio (13.4% has high *RLO*, 13.0% has low *RLO*).

[21] Moreover, the breakdown with respect to both locality and S&P 500 status changes along the same lines: From 18%, 14%, 42%, and 26% portfolio shares invested into local S&P 500, local non-S&P 500, nonlocal S&P 500, and nonlocal non-S&P 500 stocks at the 250-mile cutoff, respectively, to the respective portfolio shares of 11%, 10%, 49%, and 30% at the 100-km cutoff.

Finally, the 100-km cutoff allows for some direct comparisons between the results pertaining to individual investors and those pertaining to mutual fund managers (Coval and Moskowitz (2001)). In terms of portfolio locality shares, the average fraction of the mutual fund portfolio that is local is only 7%, whereas the average household portfolio share of local investment in our sample is 21%. In terms of performance of local investments, Coval and Moskowitz find that the difference in raw returns from investing locally versus nonlocally is 2.7% over the 1-year horizon (1.2% difference in risk-adjusted returns). Our estimates, backed out from the unreported analyses using the 100-km cutoff, suggest that the difference in raw returns from investing locally is 3.8% for individual investors (1.2% difference in risk-adjusted returns).

This result should be interpreted with due caution. It does not necessarily suggest that individual investors in some sense have outperformed the professional money managers. Rather, it suggests that there indeed is locally available value-relevant information to be discovered, and that the primary source of mutual fund managers' success along this dimension likely is their awareness of the existence of such information and their diligence in collecting it, rather than some particularly insightful and sophisticated way of interpreting it once it is available to them.

V. Conclusion

This study provides a detailed insight into the availability of asymmetric information in financial markets, particularly in the context of geography of investment by individual investors. Using a detailed data set on the investments that 78,000 U.S. retail investors made through a large discount broker over the 6-year period from 1991 to 1996, we find that individual investors exhibit local bias, that is, disproportionate preference for local stocks, to an even larger degree than U.S. mutual fund managers do (see Coval and Moskowitz (1999)).

As for the preference for local stocks, we sought to discern whether such local bias is primarily a result of the individual investors' ability to exploit asymmetric information or their inclination simply to invest in the companies they are familiar with (though not necessarily particularly informed about). We find remarkably robust evidence of superiority of individual investors' local investments relative to their nonlocal investments. This evidence is particularly strong where information asymmetries are likely to play the most pronounced role—among the non-S&P 500, less widely known stocks.

The significance of the results presented herein stretches beyond the concerns related to individual investors; the results reported in this paper help us refine our understanding of the professional managers' ability to exploit local information (Coval and Moskowitz (2001)). Specifically, two very different groups of investors posted similar returns to locality (differential performance between local and nonlocal stocks). Thus, both individual households and professional money managers seem to have been able to process and exploit locally available information to earn excess returns.

REFERENCES

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

Benartzi, Shlomo, 2001, Excessive extrapolation and the allocation of 401(k) accounts to company stock, *Journal of Finance* 56, 1747–1764.

Bodie, Zvi, and Dwight B. Crane, 1997, Personal investing: Advice, theory, and evidence, *Financial Analysts Journal* 53, 13–23.

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.

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

Cooper, Ian, and Evi Kaplanis, 1994, Home bias in equity portfolios, inflation hedging, and international capital market equilibrium, *Review of Financial Studies* 7, 45–60.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *Journal of Finance* 54, 1–39.

Coval, Joshua D., and Tobias J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

French, Kenneth R., and James M. Poterba, 1991, Investor diversification and international equity markets, *American Economic Review* 81, 222–226.

Gehrig, Thomas, 1993, An information based explanation of the domestic bias in international equity investment, *Scandinavian Journal of Economics* 95, 97–109.

Goetzmann, William N., and Kumar, Alok, 2001, Equity portfolio diversification, NBER Working paper no. 8686.

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

Huberman, Gur, 2001, Familiarity breeds investment, *Review of Financial Studies* 14, 659–680.

Kennickell, Arthur B., and Martha Starr-McCluer, 1994, Changes in family finances from 1989 to 1992: Evidence from the Survey of Consumer Finances, *Federal Reserve Bulletin* 80, 861–882.

Liang, Nellie, and Scott Weisbenner, 2002, Investor behavior and the purchase of company stock in 401(k) plans—The importance of plan design, NBER Working paper no. 9131.

Tesar, Linda L., and Ingrid M. Werner, 1997, Home bias and high turnover, *Journal of International Money and Finance* 14, 467–492.

U.S. Census Bureau. *Gazetteer Place and Zip Code Database*. Available at http://www.census.gov/geo/www/gazetteer/places.html.

Weisbenner, Scott, 2002, Do pension plans with participant choice teach households to hold more equity?, *Journal of Pension Economics and Finance* 1, 223–248.

Zhu, Ning, 2003, The local bias of individual investors, Working paper, Yale School of Management.