Individual Investors, Location, and Portfolio Choice

Mark S. Seasholes
HKUST

Mingzhu Tai
Tsinghua University

Zhishu Yang
Tsinghua University

This Version: 28-Dec-2010*

Abstract

We exploit a unique feature of Chinese stock markets to test relations between an individual investor's portfolio holdings and location. Investors tilt their holdings towards locally headquartered companies (+5.95% more than the market-cap weighted portfolio), companies from similar cultural backgrounds (+5.73%), and companies listed near where the investors lives (+10.80%). These three effects are distinct and co-exist. There is no evidence of investors having value-relevant information about stocks they hold. We do, however, provide evidence of long-lived familiarity effects. We use the location of IPO managers as instruments for increased familiarity with local companies (many IPOs took place 10+ years ago and the location of a manager is not linked with a company's current economic environment.) Increased familiarity is associated with portfolio tilting towards local companies that is 1.22× to 1.44× it otherwise would have been.

Keywords: Portfolio Choice, Location, Home Bias
JEL Number: G11, G15, D1

*We thank the original authors of Feng and Seasholes (2004b) for sharing their data, methodologies, and results. This paper expands the earlier, unpublished work. The earlier paper benefited from seminar discussions at UC Berkeley, Duke, HKUST, the research department at BGI, the PBFA Conference, and the NBER Asset Pricing meetings. Contact information: **Mark S. Seasholes**, HKUST Dept of Finance (Rm 2413), Clear Water Bay, Kowloon Hong Kong, HK Tel: +852.2358.7668, USA Tel: +1.510.931.7531, Email: Mark.Seasholes@gmail.com; **Mingzhu Tai**, Email: taimzh.05@sem.tsinghua.edu.cn; **Zhishu Yang**, Email: yangzhsh@sem.tsinghua.edu.cn; Both Tai and Yang are at the Department of Finance, School of Economics and Management, Tsinghua University, Beijing, 100084 P.R. China, PRC Tel: +86.10.6277.1769, Fax: +86.10.6278.4554.

Individual Investors, Location, and Portfolio Choice

This Version: 28-Dec-2010

Abstract

We exploit a unique feature of Chinese stock markets to test relations between an individual investor's portfolio holdings and location. Investors tilt their holdings towards locally headquartered companies (+5.95% more than the market-cap weighted portfolio), companies from similar cultural backgrounds (+5.73%), and companies listed near where the investors lives (+10.80%). These three effects are distinct and co-exist. There is no evidence of investors having value-relevant information about stocks they hold. We do, however, provide evidence of long-lived familiarity effects. We use the location of IPO managers as instruments for increased familiarity with local companies (many IPOs took place 10+ years ago and the location of a manager is not linked with a company's current economic environment.) Increased familiarity is associated with portfolio tilting towards local companies that is 1.22× to 1.44× it otherwise would have been.

Keywords: Portfolio Choice, Location, Home Bias
JEL Number: G11, G15, D1

1 Introduction

It is well known that investors tilt their equity portfolios towards local companies. In fact, a search for the term "home bias" on EconLit, JSTOR, and SSRN yields over 1,300 citations. About half of the papers were written in the last five years. In addition to showing a preference for local companies, investors exhibit a "cultural affinity" towards companies in which they and the CEOs share the same language and culture. Finally, a company's "location of trade" (i.e., where the stock is listed) appears to influence who holds shares.[1]

Despite the wealth of research studying home bias, cultural affinity, and location of trade, it remains difficult to disentangle these three effects. Do investors living in Sydney show a preference for shares of BHP Billiton Limited because the company is headquartered in Melbourne? Or, because part of the company used to be Broken Hill Proprietary Company (an Australian firm)? Or, because shares of BHP trade on the Australian Stock Exchange? As we can see from the BHP example, disentangling the three effects is difficult.

In this paper, we exploit institutional features of the Chinese capital markets. First, investors who open an account at a given brokerage branch office must place all trades through that office. Since telephone and internet trading was rare in the early 2000s, this rule meant that investors opened accounts at offices near where they lived. Once we know where an investor lives, we can test if investors have a preference for locally-headquartered companies.

Second, investors in China use their internal identity card to open accounts. Their identity numbers are coded with information about birthdate, gender, and city where the person is registered. We focus on differences between this city and the location of the brokerage office where the investor trades. Finding a difference allows us to identify investors who have moved from their birthplace. Once we have identified investors who have moved, we use the fact that China is a vast country with many local dialects. We test if investors have a preference for companies located in their birth regions.

Third, China has two stock exchanges. Companies must list on one or the other but cannot

[1] On 26-Dec-2010, we typed the term "home bias" into the three electronic databases. EconLit came up with 460 citations (273 in since 2005). JSTOR came up with 409 citations (160). SSRN had 444 citations. Grinblatt and Keloharju (2001) show that "investors are more likely to hold, buy, and sell the stocks of Finnish firms that are located close to the investor, that communicate in the investor's native tongue, and the have a chief executives of the same cultural background." Two papers that study location of trade effects are Froot and Dabora (1999) and Chan, Hameed, and Lau (2003).

list on both. Therefore, we can test whether investors living near one of the exchanges have preferences for locally listed companies (regardless of where a company's headquarters may be).

The unique features may best be understood by way of example. Consider four young men, two born in Sichuan province and two born in Shanxi province.[2] All four leave home for job-related reasons. Two men (one from each province) settle in Shanghai and the other two settle in Shenzhen in Guandong province. Each opens a brokerage account in his new location. We can test whether any of these investors has a preference for Shanghai-headquartered or Guangdong-headquartered stocks. We can also test whether an investor from Sichuan prefers Sichuan-headquartered companies while an investor from Shanxi prefers companies from his home province.

We expand the above example by considering stocks of two companies that are both headquartered in an unrelated province (say Yunnan.) If one company is listed in Shanghai and the other in Shenzhen, we can test for location of trade effects. The holdings of our four young men, specifically because they are not from Yunnan, provide power to disentangle location of trade effects from home bias and cultural affinity.

The first major result of our paper is that all three effects coexist. Investors tilt their holdings towards locally headquartered companies (+5.95% more than the market-cap weighted portfolio), companies from similar cultural backgrounds (+5.73%), and companies listed near where the investors lives (+10.8%). These three effects are distinct. When we simultaneously test for all three in a multi-variate setting, no one effect "drives out" any other effect.

The second major result of our paper is that, economically, the location of trade effect appears particularly strong. In terms of tilting an investor's portfolio away from the market portfolio, we estimate the home bias effect is 3.31× stronger than the cultural affinity effect. However, the location of trade effect is 1.73× stronger than the home bias effect. The estimate and compare the economic magnitudes of these three effects is only possible due to the institutional set-up in China.

We test whether the portfolio choices of investors in our sample appear based on value-relevant information. We find no evidence of this. Investors do not earn abnormal returns on their local holdings. In fact, one's local holdings do not out-perform one's remote holdings.

[2] All provinces and regions are shown on the map of the PRC provided in Figure 1.

The third major result of our paper is that holdings reflect long-lived familiarity with a stock. We use the location of IPO managers as instruments for increased familiarity with a company (many IPOs took place 10+ years ago and the location of a manager is not linked with a company's current economic environment.) Increased familiarity is associated with portfolio tilting towards local companies that is 1.22× to 1.44× normal levels.

The results in this paper are based on extensive brokerage data. Our holdings data start in 2000 and end in 2009. Overall, there are 461,067 individuals and we have personal information on 377,669 of them. Branch offices are located in 17 of 31 regions in the PRC. Investors were born, and companies are located, in all 31 regions. On a single date (31-Dec-2007) investors in our data hold approximately RMB 40 billion of stock which is approximately USD 6 billion. Studying an extensive dataset helps to better understand individual investor portfolios.

In the year 2008, Chinese (PRC) stock markets became the second largest in the world when ranked by equity market capitalization. PRC stock markets fell sharply during the financial tsunami, but again passed Japan in mid-2009. Our paper contributes to better understanding of the world's second largest equity market.

The remainder of our paper is structured as follows. Section 2 describes our data and reports overview statistics. Section 3 tests for different types of location-based biases. Section 4 tests whether investors have value-relevant information about local stocks. Section 5 tests whether a stock's IPO process has long-term effects on holdings. Section 6 concludes.

2 Data and Overview Statistics

We study a new, large, and comprehensive set of brokerage records from the People's Republic of China (PRC). The records come from a single brokerage company with multiple branch offices around the country. Our records come from 39 branch offices located in 17 different regions across the PRC. A "region" can be either a province (e.g., Fujian), municipality (e.g., Shanghai), or autonomous region (e.g., Xinjiang). Table 1 provides definitions of variables used in this paper. Figure 1 provides a map of China along with the locations of different regions. Table 2, Panel A shows the location of the branch offices in our dataset. The third column of the table shows the date of the first trading record which indicates when at least

one of the branch offices in a region was open for business. The first date in our records is 04-Jan-2000 and the last date is 31-Dec-2009.

[Insert Table 1, Figure 1, and Table 2]

In the PRC, individuals open brokerage accounts using internal identity cards. Thus, a single "brokerage account" can be linked to a single person. It is possible for one individual to control multiple brokerage accounts by gathering identity cards from neighbors and opening brokerage accounts in their names. To control for this possibility, we consider only "fund accounts" which are internal codes that link a single individual to one or more brokerage accounts. Throughout this paper, our unit of an analysis is a single individual who may control one or more underlying brokerage accounts.

The fourth column shows the total number of individuals who traded at some point during our sample period. Overall, we have data from 461,067 individuals for an average of 27,122 individuals per region. The highest number of individuals are from Liaoning since the brokerage company originally started its business in the province.

Identification of an individual's current location is based on the location of the branch office. In the PRC, one must place trades through the branch office in which he originally opened the account. This rule is different from the USA (for example) where a Schwab customer from San Francisco can place trades at any Schwab office. In the early years of stock trading in the PRC, internet and phone transactions were rare. This meant that individuals had to physically travel to the local brokerage office each time they wanted to trade. Feng and Seasholes (004a) exploit this PRC market structure when studying trading patterns in the PRC.

The internal PRC identity number is encoded with an individual's date of birth, gender, and "Hukou" location. The "Hukou" system is a method of registering people to live in certain regions. If the Hukou location is the same as the brokerage office location, either the investor was born in the region or the investor officially moved to the region. If the Hukou location is different from the brokerage office location, one can assume the investor has moved and is not officially registered in the new location.

Table A, Column 5 shows that we have birth data for 377,669 investors. The 83,398 individuals without birth data likely registered their accounts without providing their internal identity number (and thus broke a rule). Sometimes brokers are rushed or lazy and do not enter all the required information.

[Insert Figure 2A]

Figure 2A shows the evolution of our dataset over time. On 31-Dec-2000, there are only five brokerage offices and a total of 15,000 individuals. There is a large increase by 2003 to 30 brokerage offices and 127,000 individuals. Therefore, when we use panel data, we mainly consider 2003 to 2009. By 2009 there are 42 brokerage offices and 320,000 individuals.

2.1 Portfolio Overview Statistics

Table 2, Panel B shows holdings data on a specific date (31-Dec-2007). On this date, our dataset contains 266,317 individuals with holdings in 1,439 different stocks. The overall median number of stocks held by a given individual is two. The average individual's portfolio value varies between RMB 99,878 in Tianjin to RMB 687,550 in Zhejiang. This wide range is emblematic of wealth dispersion in the PRC. Costal areas such as Zhejiang, Shanghai, and Guangdong are considerably more wealthy than interior areas such as Hebei, Hunan, and Sichuan. The approximate exchange rate is RMB 6.68 per USD meaning that RMB 687,550 is approximately USD 100,000.

Portfolio values are positively skewed and median portfolio values range from RMB 28,730 (Hebei) to RMB 78,528 (Shanghai). The overall median is RMB 38,200 on 31-Dec-2007. The final column in Panel B shows the total value held by all investors. Our dataset is comprised of RMB 39.5 billion on 31-Dec-2007 which works out to just a little under USD 6 billion.

[Insert Figure 2B]

Figure 2B shows the evolution of total portfolio holdings over time. On 31-Dec-2000, there is only RMB 1.5 billion held in stocks. By 2003, there is a large jump to RMB 6.7 billion. The growth of the Chinese stock markets and the result of the financial Tsunami are clearly

visible in the figure. Total holdings reach RMB 39 billion in 2007, fall to RMB 18.5 billion in 2008, and then rebound to RMB 44.8 billion by 2009.

2.2 Additional Datasets

In addition to the brokerage account data, we obtain closing stock prices and shares outstanding from the CSMAR database. We use the prices to value investors' portfolios at the end of each year. We also use the prices to value weekly calendar-time portfolios. Shares outstanding and prices are used to calculate market capitalizations.

There are 1,527 listed stocks (A-shares only) as of 31-Dec-2007. Of these stocks, 850 are listed on the Shanghai Stock Exchange and 677 are listed on the Shenzhen Stock Exchange in Guangdong province. On same date, investors in our dataset hold 1,439 of the 1,527 listed stocks. Of these, 815 are listed in Shanghai and 624 stocks are listed in Shenzhen. Appendix A provides a geographic overview of Chinese stocks.

We obtain details related to initial public offerings (IPOs) in the PRC from the Resset database. In total, we have IPO manager data for 1,912 stocks, some of which have been delisted. In total, investors in our dataset hold 1,439 different stocks on 31-Dec-2007. Of these, we have IPO manager data for all 1,439 stocks. Note that not all of the 1,439 stocks are held by a local investor. As far as timing, 10.9% of IPOs took place in 1993 or earlier, 39.5% of IPOs took place between 1994 and 1998, 28.6% took place between 1999 and 2003, the remaining 21% took place in 2004 or later.

3 Location Biases and Portfolio Tilting

3.1 Locally-Headquartered Companies

We test whether an individual investor prefers stocks of locally-headquartered companies. An investor's location is defined by the region of the brokerage office where he opened his account and trades. A company's location is defined by the region where its headquarters are located. For each investor, we measure the fraction of his portfolio that is invested in companies whose headquarters are in the same region where the investor currently lives. We

then report the average fraction for all investors from a given region. We also calculate the fraction of the total market capitalization that is headquartered in the same region.

Table 3A reports results from a single date–31-Dec-2007. The table reports three measures of overweighting (also known as portfolio tilting). The first measure represents the difference between Column (3) and Column (4). A positive difference indicates that investors overweight locally-headquartered stocks. We see the difference measure in Column (5) is positive for all regions. The difference is greater than 10% in Jilin, Shanghai, and Tianjin. Taking a weighted average of this overweight measure, we see investors overweight local stocks by 5.95%. The weights are based on the number of individuals shown in Column (2).

To control for the fact that some regions have higher fractions of market capitalization than other regions, Column (6) reports a second overweighting measure. Measure #2 is defined as a ratio equal to Column (3) ÷ Column (4) minus one. The weighted average of this measure is 2.67 indicated that investors place 2.67× more weight in local stocks than a market-cap weighted portfolio predicts. Five regions have measures over four including Jilin, Neimenggu, Tianjin, Xinjiang, and Yunnan. All five of these regions are on one of the PRC's national borders.

Column (7) reports a third measure defined as the natural log of (3) ÷ (4). The weighted average is 1.18 and well above zero. Six regions have log ratios greater than 1.5 including the five mentioned in the paragraph above plus Shaanxi.

[Insert Figure 3A]

Figure 3A shows the evolution of our the three measures over time. Measure #1 is highest in 2000, 2001, and 2002 when there are only a few, concentrated brokerage offices in our dataset. The measure is basically flat at 6% except for a slight bump in 2006. Measures #2 and #3 show upward trends from 2003 to 2009. Measure #2 is 1.92 in 2003 and 3.63 in 2009. Measure #3 is 0.97 in 2003 and 1.43 in 2009. All measures are positive (at all times) indicating tilting one's portfolio towards local stocks is pervasive.

We end this section by concluding that individuals in the PRC significantly overweight locally-headquartered stocks.

3.2 Cultural Affinity

We test whether an individual investor prefers companies from a similar cultural background as the individual. Regions in the PRC typically have distinct dialects, food, and customs. Dialects are typically not understood from one region to the next and Mandarin is taught as a unifying language. An investor's cultural background is estimated by looking at coding in his Hukou number. If the investor moves and does not re-register with the Hukou system, his Hukou number reflects where he was born. If the investor has re-registered, he is issued a new Hukou number that reflects where he currently lives. Thus, our noisy estimate of an investor's cultural background, biases results against finding any correlation between cultural background and holdings. A company's location is defined by the region where its headquarters are located.

For each investor, we measure the fraction of his portfolio that is invested in companies whose headquarters are located in the same region where the investor was born. We then report the average fraction for all investors from a given region. We also calculate the fraction of the total market capitalization that is headquartered in the same region.

Table 3B reports results from a single date–31-Dec-2007 and follows the same format with three measures of overweighting (also known as portfolio tilting). The overweighting difference measure in Column (5) is positive for all regions. The difference is greater than 10% in Jilin and Shanghai. Taking a weighted average of this overweight measure, we see investors overweight local stocks by 5.73%. The weights are based on the number of individuals shown in Column (2).

Measure #2 is a ratio and its weighted average is 2.57 indicating that investors place 2.57× more weight in local stocks than a market-cap weighted portfolio predicts. Three regions have measures over four including Jilin, Xinjiang, and Yunnan. Measure #3 is the natural log of a ratio. The weighted average is 1.15 and well above zero.

We end this section by concluding that an individual in the PRC significantly overweights companies headquartered in regions matching the investor's cultural background. Later, we test whether this preference is distinctly different from simply liking to hold locally-headquartered stocks.

3.3 Location of Trade

We test whether an individual has a preference for companies listed near where the investor currently lives. Again, investor location is defined by the region of the brokerage office where he opened the account and trades. Companies can only list on one of the PRC's two stock exchanges. One exchange is located in Shanghai and the other is in Shenzhen, Guangdong Province. We focus on investors from these two regions and consider holdings of all stocks.

For a Shanghai-based investor, we measure the fraction of his portfolio that is invested in Shanghai-listed companies. We then subtract the fraction of Shanghai-listed companies in the PRC and report the difference. We repeat this procedure for the three other combinations of investor location and exchange location—Guangdong investors and Shanghai-listed stocks; Shanghai investors and Guangdong stocks; and Guangdong-investors and Shanghai stocks.

Table 3C reports results from as of 31-Dec starting in 2003 and ending in 2009. Panel A considers only Shanghai-listed stocks. We see Shanghai-based investors overweight these stocks by 4% to 14%. There is a generally declining trend over time. Guangdong-based investors underweight the same stocks by 5% to 20%.

Panel B considers only Guangdong-listed stocks and is a mirror image of the panel above. We see Guangdong-based investors overweight these stocks by 5% to 20%. Shanghai-based investors underweight the same stocks by 4% to 14%.

We end this section by concluding that individuals in the PRC significantly overweight companies that are listed near where the investor lives.

3.4 Regression Analysis of Portfolio Tilting

We test whether an individual investor prefers locally-headquartered stocks, culturally-similar companies, and/or locally-traded companies. To carry out the test, we follow these steps: 1) We choose 31-Dec-2007 as a single date. Using additional dates increases the statistical power of our tests which is not a problem here. 2) We calculate each investor's weight in every stock in the market. Given that the median portfolio has only two stocks, most of these weights are zero. We worry about overstating statistical significance due to all the zero-observations and make corrections later. 3) We calculate the weight of each stock

in the market. 4) We calculate the difference between each investor-stock weight and the market weight. 5) We then regress the over/under weight measures on a series of indicator variables.

$$\omega_{i,j} - \omega_{m,j} \quad = \quad \beta_1 D_{i,j}^{R(i)=HQ(j)} + \beta_2 D_{i,j}^{B(i)=HQ(j)} + \beta_3 D_{i,j}^{R(i)=LT(j)} + \epsilon_{i,j} \tag{1}$$

Above, $\omega_{i,j}$ is weight of stock j in investor i's portfolio and $\omega_{m,j}$ is the weight of stock j in the market portfolio. The indicator variable $D_{i,j}^{R(i)=HQ(j)}$ takes a value of one if investor i lives in the same region as stock j's headquarters. The indicator variable $D_{i,j}^{B(i)=HQ(j)}$ takes a value of one if investor i was born in the same region as stock j's headquarters. The indicator variable $D_{i,j}^{R(i)=LT(j)}$ takes a value of one if investor i lives in the same region as where stock j's is traded.

[Insert Table 4]

Table 4 shows the regression results. Regressions (1), (2), and (3) mirror earlier findings. Investors overweight locally headquarterd stocks, stocks headquartered where the investors were born, and stocks that trade on nearby exchanges. Regression (4) is the main result. We see all three effects co-exist. No single effect "drives out" another. We can estimate the relative strength of each effect by looking at the change of coefficients from the first three regressions to Regression (4). The first effect remains strong as β_1 only falls from 2.91 to 2.25. Cultural affinity is probably the weakest effect as β_2 falls from 2.84 to 0.68. The location of trade effect is quite strong as β_3 only falls from 0.36 to 0.26.

As noted on the table, there are 14,741,554 total observations. Of these, $\omega_{i,j}$ is non-zero for 728,175 while the vast majority of observations are "zero"-holdings. To calculate statistical significance, we consider the number of independent observations to be only 728,175. This is conservative as investors could have actively chosen to assign zero weight to certain stocks. Regardless, results are statistically significant at all conventional levels.

Table 4 contains this paper's first major result. The institutional set-up in the PRC allows us to study portfolio holdings along three dimensions. These dimensions co-exist and no single effect drives out one of the others.

3.5 Economic Significance

We estimate economic significance using Table 4, Regression (4). This regression is estimated with the unit of analysis being an investor-stock holding. There are about 1,500 listed stocks on 31-Dec-2007. Dividing by 30 gives an approximate figure of 50 listed companies per region. We can think of the 2.25% coefficient on $D_{i,j}^{R(i)=HQ(j)}$ as affecting approximately 50 stocks.

A similar calculation as the one described above can be used for the $D_{i,j}^{B(i)=HQ(j)}$ coefficient. The the first two coefficients are relevant for the same number of stocks (roughly). If we divide 2.25 by 0.68 we can say that the home bias effect is roughly 3.31× the magnitude of the cultural affinity effect.

The location of trade coefficient $D_{i,j}^{R(i)=Lt(j)}$ is approximate one eighth as large as the the first coefficient (0.26 vs. 2.25). For an investor living in Shanghai, however, there are more than 750 locally listed stocks. This means the coefficient on $D_{i,j}^{R(i)=Lt(j)}$ is relevant for 750/50 or about 15× more stocks than the other two coefficients. If we multiply the 0.26 coefficient by 15, we get 3.90 meaning that location of trade effects are roughly 1.73× the magnitude of the home bias effect (3.90 / 2.25 = 1.73)

The finding that the location of trade effect has the highest level of economic significance represents the second major result of our paper.

4 Performance-Based Results

We test whether investors have value relevant information about the stocks they hold. The tests involve making holdings-based calendar time portfolios at a weekly frequency. We start with investors from a given region. We make one portfolio based on investors' local holdings where local is defined as companies headquartered in the same region as where the investor currently lives (r_t^{local}). We make a second portfolio based on investors' remote holdings (r_t^{remote}). We next calculate the difference in returns for investors in the given region as $r_t^{local} - r_t^{remote}$. We end by repeating these steps for each region.

The procedure above results in a single time series of returns, 476 weeks in length, for each region in our database. We regress the difference of returns on a constant, the riskfree rate,

and the market return. The regression takes the form of:

$$r^{local}_{reg,t} - r^{remote}_{reg,t} = \alpha + \gamma_1 r_{f,t} + \gamma_2 r_{m,t} + \epsilon_{r,t} \quad (2)$$

Above, $r^{local}_{reg,t}$ is the weekly return of local portfolio from investors in region "reg" and $r^{remote}_{reg,t}$ is the weekly return of their remote portfolio.

Table 5 shows the results of our calendar-time holdings based portfolios. Panel A considers the full dataset of 476 weeks. In Regression (1), we see that local holdings outperform remote holdings by 1.82 basis points per week but the difference is not statistically significant. In Regression (2) we see the alpha is -2.47 basis points per week and not statistically significant.

[Insert Table 5]

Table 5, Panel B uses only data from 2003 onward for a total of 330 weeks. Regression (1) shows an alpha of 2.17 basis points per week with no statistical significance. Regression (2) shows an alpha of -1.94 basis points per week with no statistical significance.

We conclude that there is no evidence of investors having value-relevant information about the local stocks they hold. Our results are based on holdings-based calendar-time portfolios. Our regressions do not lack power as they are based on 5,000+ observations.

5 Long-Lived Familiarity Effects

We test whether events in the past affect investors' holdings in the present. In particular, we are interested in events that are unlikely to give investors relevant information about a company's current outlook. Our tests involve studying the location of the investment bank that ran a given company's initial public offering. We test whether investors who were more exposed to marketing materials about a company during its IPO process, are more likely to hold the company's stock today.

Our test involve cross-sectional regressions at a single point in time (31-Dec-2007). On that date, investors in our dataset held 1,439 different stocks. For each of these 1,439 companies, we calculate the total number of shares held by all investors in our sample. For each company,

we then count the fraction of shares held by investors who currently live in the same region as the companys headquarters. We call this our aggregate measure of location-based portfolio tilting (Ω_j) for stock j. It is important to note that we do not have the complete holdings of each listed-company. Therefore, Ω_j contains sampling noise which should bias results against us.

We consider only the 1,056 companies where we have at least one investor in our sample, who lives in the same region as the firms headquarters, and holds some of the firms stock. One can think of our test as follows: we take home bias as given and then test for cross-sectional differences (if any) in a home bias measure across firms. For each of the 1,056 companies, we regress our aggregate measure on a constant, an indicator variable (D_j^{Ibank}), and the company's market capitalization. The indicator variable, D_j^{Ibank}, takes a value of one if the firm is headquartered in the same region as the investment bank that managed the IPO. This is a single cross-sectional regression (across stocks) that uses holdings as of 31-Dec-2007. The regression specification is:

$$\Omega_j \quad = \quad \alpha + \beta_1 D_j^{Ibank} + \beta_2 ln(MktCap) + \beta_3 D_j^{Ibank} \cdot ln(MktCap) + \epsilon_j \tag{3}$$

[Insert Table 6]

Table 6 reports the main results. in Regression (1) we see the fraction of shares held in home region-firms is 1.22× normal levels due to the investment bank being headquartered in the same region. To see this 1.22× difference, notice that local holdings are 10.10% of shares outstanding for companies where the investment bank was not local and are 12.35% (=0.1010+0.0225) for companies where the investment bank was local.

Regression (2) interacts the investment banking dummy with year indicators. Most of the interactive coefficients remain near 2.25% except the 1999 to 2001 range. We have the highest level of significance from older IPOs that took place in 1998 or earlier. Finally, in Regression (2) we include the natural log of a firm's market capitalization. The marginal effect of having a local investment bank appears to associated with local holdings being 1.44× higher (going from 4.76% to 6.86%).

We conclude that investor holdings appear influenced by events and factors that are not associated with today's economic environment. For example, investors have a preference

for locally-headquartered companies. This preference increases by 1.22× to 1.44× if a local company was taken public by a local investment bank. Most surprisingly, our results hold *even if the IPO took place ten or more years ago!*

6 Conclusions

This paper studies individual investors' portfolio choices with a focus on geography. We use unique institutional features of the Chinese stock markets to disentangle three effects. First, investors overweight locally-headquartered stocks (pure home bias). Second, investors overweight stocks headquartered from the region of the country where the investor was born (cultural affinity). Third, investors overweight stocks listed on nearby exchanges (location of trade bias).

We find the three effects mentioned above coexist. No one effect drives out either of the other two. In terms of economic significance, the location of trade bias is particularly important in explaining shifts away from the passive (market-cap weighted) benchmark. The is no evidence that tilting one's portfolio away from the market portfolio is driven by individuals having value-relevant information.

We do, however, find evidence of long-lived familiarity effects. We use the location of the securities firm that managed the IPO as an instrument for increased levels of investor familiarity of with a stock. While local securities firms can market new issues near their headquarters, it is hard to argue that information released during the IPO process has information about a company's prospects years later.

Our results indicate that increased familiarity with a stock, even *ten years prior* is associated with portfolio tilting towards local companies that is 1.22× to 1.44× normal. This final result of our paper is particularly fascinating and suggests a number of avenues for future work. How do market efforts influence portfolio choice? How does the influence vary across investor age? Experience? Wealth? Can long-lived familiarity effects help financial economists estimate the magnitude of search costs faced by individual investors. Attacking these questions is left for future work.

References

Chan, Kalok, Allaudeen Hameed, and Sie Ting Lau. 2003. "What If Trading Location is Different from Business Location? Evidence from the Jardine Group." *Journal of Finance* 58(3):1221–1246.

Feng, Lei, and Mark S. Seasholes. 2004a. "Correlated Trading and Location." *Journal of Finance* 54(5):2117–2144.

Froot, Kenneth A., and Emil M. Dabora. 1999. "How Are Stock Prices Affected by the Location of Trade?" *Journal of Financial Economics* 53:189–216.

Grinblatt, Mark, and Matti Keloharju. 2001. "How Distance, Language, and Culture Influence Stockholdings and Trades." *Journal of Finance* 56(3):1053–1073.

A Geographic Distribution of Listed Stocks

We provide the geographic distribution of listed stocks as of 31-Dec-2007. The table focuses on the same 17 regions where our brokerage offices are located. In total, there are 1,527 listed stocks of which investors in our data hold 1,439 different ones.

Location of Branch Offices (Region)	Number of Listed Stocks (Num)	Number of Shanghai Stocks (Num)	Number of Guangdong Stocks (Num)	Market Cap of Listed Stocks (RMB bn)
Fujian	54	32	22	171.6
Guangdong	181	35	146	1,574.6
Hebei	37	19	18	190.1
Heilongjiang	28	21	7	89.9
Hunan	47	18	29	173.2
Jiangsu	110	65	45	382.9
Jilin	34	20	14	144.0
Liaoning	51	26	25	266.4
Neimenggu	21	15	6	105.0
Shaanxi	26	13	13	67.1
Shandong	85	47	38	378.5
Shanghai	144	135	9	1,198.6
Sichuan	66	35	31	316.6
Tianjin	27	18	9	234.1
Xinjiang	30	21	9	132.8
Yunnan	26	13	13	123.7
Zhejiang	111	59	52	314.2
Other	449	258	191	3,220.0
Total	1,527	850	677	9,083.3

Table 1: Variable Descriptions and Definitions

Branch Office	Our dataset contains 39 branch offices of the same securities firm. The branch offices are located in 17 of 31 regions in the PRC.
Company Location	Based on the company's headquarters. Companies are headquartered in 31 of 31 regions in the PRC.
Cultural Background	Based on region identified in investor's Hukou number. Investors in our dataset come from 31 of 31 regions in the PRC.
Fund Account	One or more brokerage accounts controlled by a single individual.
Hukou System	System of registering people to live in areas of the PRC.
Individual Investor	An individual in the PRC that is identified by a unique Hukou number and/or unique Fund Account identifier.
Investor Location	Based on where the investor opened his brokerage account and where he places trades. Investors in our dataset are located in 17 of 31 regions in the PRC.
Region	Can be either a province (e.g., Fujian), municipality (e.g., Shanghai), or autonomous region (e.g., Xinjiang). We identify 31 different regions in the PRC. Each region has a unique regional code and is shown in Figure 1.
$\omega_{i,j}$	The weight of stock j in investor i's portfolio.
$\omega_{m,j}$	The weight of stock j in the market portfolio.
Ω_j	The fraction of stock j held by investors in the same region. All figures use data from our brokerage offices only.
$D_{i,j}^{R(i)=HQ(j)}$	Indicator variable that takes a value of one if investor i lives in the same region as stock j's headquarters.
$D_{i,j}^{B(i)=HQ(j)}$	Indicator variable that takes a value of one if investor i was born in the same region as stock j's headquarters.
$D_{i,j}^{R(i)=LT(j)}$	Indicator variable that takes a value of one if investor i lives in the same region as where stock j is being traded.
$D_{i,j}^{Ibank}$	Indicator variable that takes a value of one if the investment bank that ran a company's IPO is located in the same province as where the investor currently lives.

Table 2: Summary Statistics

The table presents summary statistics of our brokerage data. The data come from 39 branch offices of a single securities firm. The offices are located in 17 different regions in the PRC. The table shows the date of the first trade in each region. We count the number of different individuals who traded at some time during our sample period. The final column shows the number of individuals with birth data. Panel B provides overviews of portfolios as of 31-Dec-2007.

Panel A: Overview Stats

Location of Branch Offices (Region)	Number of Branch Offices (Num)	Date of the First Transaction (D-M-Y)	Num of Indivs (Num)	Indivs with Birth Data (Num)
Fujian	1	23-Mar-2005	5,701	5,388
Guangdong	3	04-Jan-2000	11,046	8,326
Hebei	2	02-Jan-2003	29,930	25,938
Heilongjiang	1	02-Jan-2003	10,594	7,441
Hunan	2	02-Jan-2003	24,985	20,702
Jiangsu	2	02-Jan-2003	9,693	7,379
Jilin	3	02-Jan-2003	96,769	83,755
Liaoning	9	20-Oct-2000	155,647	129,075
Neimenggu	1	19-Feb-2003	5,151	4,431
Shaanxi	1	02-Jan-2003	4,619	4,119
Shandong	1	01-Apr-2005	2,819	2,542
Shanghai	3	02-Jan-2003	15,377	10,418
Sichuan	5	04-Jan-2000	56,347	40,579
Tianjin	1	02-Jan-2001	20,029	17,254
Xinjiang	1	02-Jan-2003	5,317	4,682
Yunnan	1	06-Jan-2003	2,230	1,548
Zhejiang	2	04-Mar-2005	4,813	4,092
Total	39	n.m	461,067	377,669
Average	n.m	n.m	27,122	22,216

Table 2: *Continued*

Panel B: Portfolio Overview Stats as of 31-Dec-2007

Location of Branch Offices (Region)	Num of Indivs (Num)	Median Num of Stocks Per Individual (Num)	Average Portfolio Value (RMB)	Median Portfolio Value (RMB)	Total Value Held by All Investors (RMB mm)
Fujian	2,405	3	141,145	34,871	338
Guangdong	6,176	3	371,481	75,924	2,279
Hebei	17,997	2	100,419	28,730	1,793
Heilongjiang	6,609	3	164,882	42,878	1,084
Hunan	15,382	2	108,817	33,250	1,657
Jiangsu	5,561	3	244,979	67,337	1,357
Jilin	49,957	2	118,833	34,980	5,893
Liaoning	95,059	2	136,879	38,570	12,902
Neimenggu	2,562	3	196,475	43,475	500
Shaanxi	2,629	3	230,335	32,500	604
Shandong	1,492	3	341,636	32,280	508
Shanghai	8,874	3	338,946	78,528	2,994
Sichuan	31,762	2	120,064	37,190	3,784
Tianjin	12,855	3	99,878	29,478	1,271
Xinjiang	3,174	2	172,193	43,838	543
Yunnan	1,339	3	237,634	52,853	317
Zhejiang	2,484	3	687,550	77,108	1,703
Total	266,317	n.m	n.a.	n.a.	39,527
Average	15,666	*2	149,488	**38,200	6,689

* *Calculated as the median number of stocks across all individuals on this date.*

** *Calculated as the median portfolio value across all individuals on this date.*

Table 3A: Company Headquarters and Portfolio Tilting

The table presents results related to portfolio tilting and the proximity of companies' headquarters. All figures are as of 31-Dec-2007. Column (3) shows the average fraction of individuals' portfolios that is invested in stocks headquarted where the investors currently live. Column (4) shows the fraction of the total market located in the same region. Columns (5), (6), and (7) show three measures of overweighting. Overweight Measure #1 is the difference (3)-(4). Overweight Measure #2 is the ratio [(3) ÷ (4)] - 1. Overweight Measues #3 is the ln[(3) ÷ (4)]. Values greater than zero indicate a preference for locally headquartered stocks.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Avg Frac of Port From Region	Avg Frac of Market	Overweight Measures		
Location of Branch Offices (Region)	Num of Indivs (Num)	Where Indiv Curr Lives (%)	Cap from Same Region (Num)	#1 Difference (Num)	#2 Ratio-1 (Num)	#3 Ln Ratio (Num)
Fujian	2,405	3.2%	1.8%	1.46%	0.82	0.60
Guangdong	6,176	18.6%	17.4%	1.16%	0.07	0.06
Hebei	17,997	4.9%	1.9%	2.97%	1.54	0.93
Heilongjiang	6,609	4.2%	1.0%	3.21%	3.36	1.47
Hunan	15,382	5.2%	1.9%	3.29%	1.74	1.01
Jiangsu	5,561	9.5%	4.1%	5.41%	1.32	0.84
Jilin	49,957	11.9%	1.6%	10.24%	6.23	1.98
Liaoning	95,059	6.8%	2.7%	4.09%	1.49	0.91
Neimenggu	2,562	6.4%	1.2%	5.18%	4.38	1.68
Shaanxi	2,629	3.4%	0.8%	2.67%	3.53	1.51
Shandong	1,492	11.9%	4.2%	7.71%	1.85	1.05
Shanghai	8,874	26.9%	14.8%	12.02%	0.81	0.59
Sichuan	31,762	9.8%	3.4%	6.42%	1.88	1.06
Tianjin	12,855	12.5%	2.5%	10.10%	4.12	1.63
Xinjiang	3,174	8.7%	1.5%	7.26%	4.97	1.79
Yunnan	1,339	9.6%	1.4%	8.22%	6.08	1.96
Zhejiang	2,484	8.7%	3.5%	5.12%	1.45	0.89
Total	266,317	n.a.	n.a.	n.a.	n.a.	n.a.
Wgt Avg	n.a.	9.1%	3.2%	5.95%	2.67	1.18
Simple Avg	15,666	9.5%	3.9%	5.68%	2.68	1.17

Table 3B: Cultural Affinity and Portfolio Tilting

The table presents results related to portfolio tilting and cultural affinity. All figures are as of 31-Dec-2007. Column (3) shows the average fraction of individuals' portfolios that is invested in stocks headquarted where the investor was born. Column (4) shows the fraction of the total market located in the same region. Columns (5), (6), and (7) show three measures of overweighting. Overweight Measure #1 is the difference (3)-(4). Overweight Measure #2 is the ratio [(3) ÷ (4)] - 1. Overweight Measues #3 is the ln[(3) ÷ (4)]. Values greater than zero indicate a preference for locally headquartered stocks.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Avg Frac. of Portfolio	Avg Frac of Market	Overweight Measures		
Location of Branch Offices (Region)	Num. of Indivs. (Num)	from Region Curr. Lives (%)	Cap. from Same Region (Num)	#1 Difference (Num)	#2 Ratio-1 (Num)	#3 Ln Ratio (Num)
Fujian	2,280	3.3%	1.9%	1.46%	0.78	0.58
Guangdong	5,219	14.3%	13.8%	0.51%	0.04	0.04
Hebei	16,144	5.0%	2.0%	2.95%	1.45	0.90
Heilongjiang	5,580	4.1%	1.0%	3.03%	2.94	1.37
Hunan	14,262	5.2%	2.0%	3.25%	1.66	0.98
Jiangsu	4,769	9.1%	4.0%	5.07%	1.26	0.82
Jilin	45,927	11.9%	1.7%	10.21%	6.11	1.96
Liaoning	86,812	6.6%	2.7%	3.88%	1.42	0.88
Neimenggu	2,312	6.3%	1.3%	4.99%	3.83	1.57
Shaanxi	2,431	3.6%	1.1%	2.51%	2.26	1.18
Shandong	1,371	11.2%	4.0%	7.11%	1.76	1.01
Shanghai	7,068	23.6%	12.8%	10.84%	0.85	0.61
Sichuan	25,927	9.6%	3.4%	6.23%	1.84	1.04
Tianjin	11,472	12.2%	2.5%	9.74%	3.93	1.60
Xinjiang	2,966	8.4%	1.6%	6.78%	4.17	1.64
Yunnan	1,093	9.8%	1.5%	8.24%	5.42	1.86
Zhejiang	2,134	8.0%	3.5%	4.53%	1.31	0.84
Total	237,767	n.a.	n.a.	n.a.	n.a.	n.a.
Wgt Avg	n.a.	8.7%	3.0%	5.73%	2.57	1.15
Simple Avg	13,986	9.0%	3.6%	5.37%	2.41	1.11

Table 3C: Location of Trade and Portfolio Tilting

The table presents results related to portfolio tilting and location of trade. All figures are as of year-end from 2003 to 2007. Rows indicate where investors currently live—we consider only investors in Shanghai or Guangdong. Panel A reports the fraction of individuals' portfolios invested in Shanghai-listed stocks less the fraction of total market cap listed in Shanghai (and adjusted to account for holdings from our particular securities firm). Details about the calculations are given in the text. Panel B reports the fraction of individuals' portfolios invested in Shenzhen-listed stocks less the fraction of total market cap listed in Shenzhen.

Panel A: Shanghai-Listed Stocks

	2003	2004	2005	2006	2007	2008	2009
Guangdong	-19.8%	-18.1%	-15.5%	-12.1%	-6.7%	-5.7%	-4.7%
Shanghai	9.5%	14.3%	14.7%	13.1%	10.8%	9.0%	4.1%

Panel B: Shenzhen-Listed (Guangdong) Stocks

	2003	2004	2005	2006	2007	2008	2009
Guangdong	19.8%	18.1%	15.5%	12.1%	6.7%	5.7%	4.7%
Shanghai	-9.5%	-14.3%	-14.7%	-13.1%	-10.8%	-9.0%	-4.1%

Table 4: Regression Analysis and Portfolio Tilting

The table presents results from multi-variate regressions. We regress the degree to which a given investor over/under-weights a stock on a series of indicator variables. $\omega_{i,j}$ is weight of stock j in investor i's portfolio and $\omega_{i,j}$ is the weight of stock j in the market portfolio. The indicator variable $D_{i,j}^{R(i)=HQ(j)}$ takes a value of one if investor i lives in the same region as stock j's headquarters. The indicator variable $D_{i,j}^{B(i)=HQ(j)}$ takes a value of one if investor i was born in the same region as stock j's headquarters. The indicator variable $D_{i,j}^{R(i)=LT(j)}$ takes a value of one if investor i lives in the same region as where stock j's is traded. Reported coefficients have been multiplied by 100. T-stats are reported below the coefficients and reflect only the number of non-zero portfolio weights.

	(1)	(2)	(3)	(4)
$D_{i,j}^{R(i)=HQ(j)}$	2.91			2.25
	(211.04)			(49.84)
$D_{i,j}^{B(i)=HQ(j)}$		2.84		0.68
		(205.61)		(15.23)
$D_{i,j}^{R(i)=LT(j)}$			0.36	0.26
			(23.18)	(17.06)
R-sqr	0.095	0.095	0.095	0.095

Each regression has 14,741,554 total observations.
Of these, $\omega_{i,j} \neq 0$ for 728,175 obs while $\omega_{i,j} = 0$ for 14,013,379 obs.

Table 5: Holdings-Based Calendar Time Portfolios

The table presents results related holdings-based calendar time portfolios. Panel A uses the full length of our data sample. Panel B starts in 2003 when the dataset contains more brokerage offices. Note, t-statistics are based on clustered standard errors where clustering is done by week. The number of clusters is equal to the number of weeks and reported below each regression.

Panel A: Full Time Series

	(1)	(2)
$\alpha(bp)$	1.82	-2.47
	(0.51)	(-1.86)
$r_{f,t}$		5.21
		(1.95)
$r_{m,t}$		0.01
		(0.49)
N(Obs)	5,867	5,782
N(Clusters)	476	471

Panel B: 2003 - 2009

	(1)	(2)
$\alpha(bp)$	2.17	-1.94
	(0.56)	(-1.41)
$r_{f,t}$		4.43
		(1.65)
$r_{m,t}$		-0.01
		(-0.82)
N(Obs)	5,277	5,277
N(Clusters)	330	330

Table 6: Long-Lived Familiarity Effects

The table presents results related to long-term memory effects and investor portfolios. We run a single, cross-sectional regression (across stocks). The left-hand side variable is the fraction of stock j's shares held by local investors measured using our dataset. The right-hand side variables include an indicator if the region where a company is headquartered is the same as the investment bank that was responsible for the initial public offering.

	(1)	(2)	(3)
Constant (%)	10.10 (23.31)	10.10 (24.47)	4.76 (1.35)
D_j^{Ibank}	2.25 (3.53)		2.10 (3.34)
$D_j^{Ibank} \cdot [2005, 2007]$		2.40 (2.06)	
$D_j^{Ibank} \cdot [2002, 2004]$		2.26 (1.92)	
$D_j^{Ibank} \cdot [1999, 2001]$		0.79 (1.21)	
$D_j^{Ibank} \cdot$(1998 or before)		3.12 (3.74)	
ln(MktCap)			0.37 (1.57)
N(Obs)	1,056	1,056	1,056

Figure 1: Map of the People's Republic of China

The figure shows location of 31 regions within the People's Republic of China (PRC). A "region" is defined as either a province (e.g., Jilin or 22), a municipality (e.g., Shanghai or 31), or an autonomous region (e.g., Xinjiang or 65).

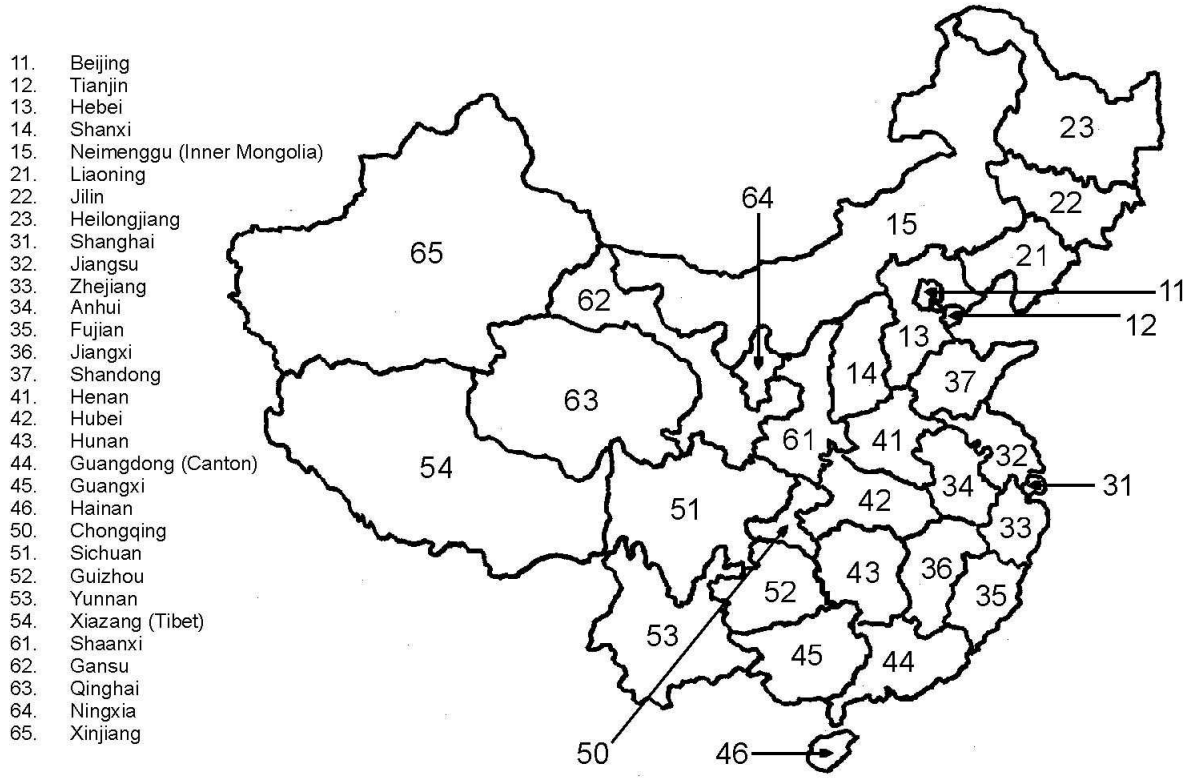

Figure 2A: Number of Branch Offices and Individuals

The figure shows the size of our dataset. The number of branch offices is shown by the columns and read on the left hand Y-axis (LHS). The number of individual investors is shown by the line and read on the right hand Y-axis (RHS).

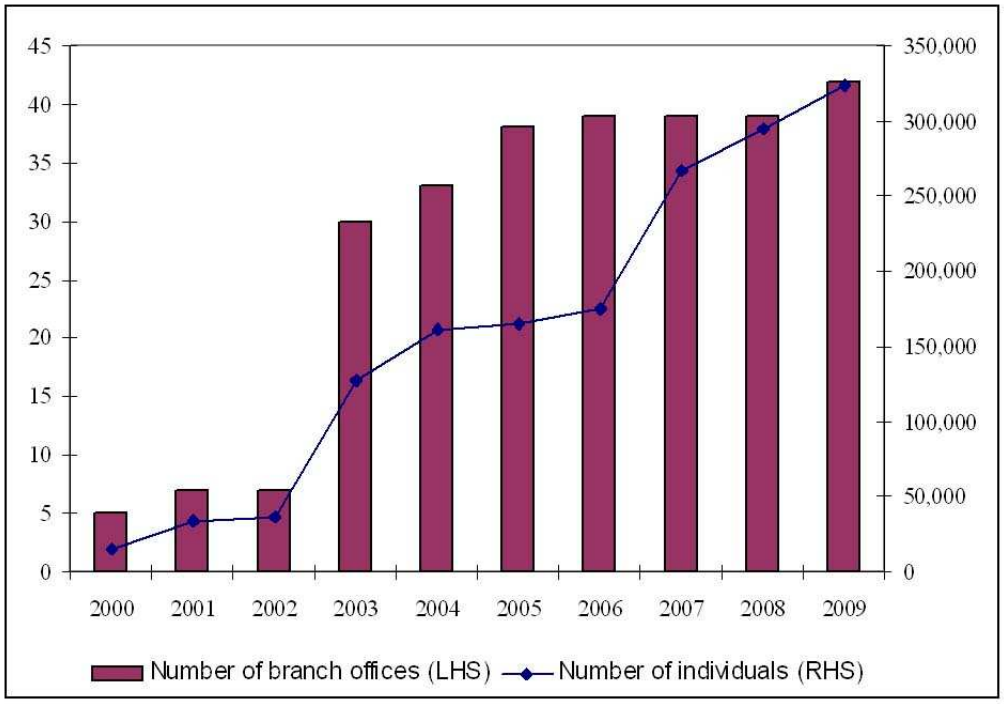

Figure 2B: Value of Holdings

The figure shows the size of our dataset measured by the value of holdings by individual investors.

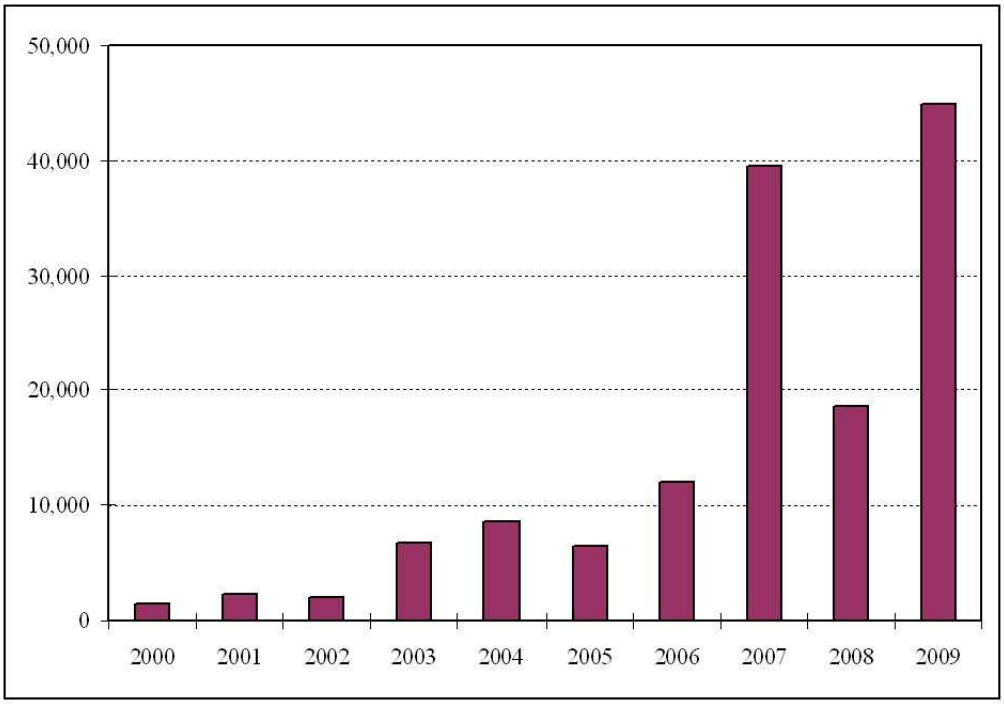

Figure 3: Company Headquarters and Portfolio Tilting

The figure shows the time series of averages related to portfolio tilting and the proximity of companies' headquarters. Consider two quantities. (A) The average fraction of individuals' portfolios that is invested in stocks headquarted where the investors currently live. (B) The fraction of the total market located in the same region. Overweight Measure #1 is the difference (A)-(B). Overweight Measure #2 is the ratio [(A) ÷ (B)] - 1. Overweight Measues #3 is the ln[(A) ÷ (B)]. Values greater than zero indicate a preference for locally headquartered stocks.