Portfolio Choice and Location of Trade

Lei Feng — McKinsey & Co.
Mark S. Seasholes — U.C. Berkeley

This Version July 12, 2004*

Abstract

This paper studies the portfolio choices of individual investors. We test the predictions of a multi-asset rational expectations model in an intra-national setting. We partition investors along geographic dimensions (provinces) and other dimensions (industries, languages, cultural groups.) These partitions exploit unique features of an emerging stock market that allow us to decompose holdings in ways that have previously not been possible. Surprisingly, stocks that are traded near where investors live receive the highest weights in their portfolios. This finding gives insight into an existing puzzle in finance regarding location of trade. Finally, we test whether portfolio choices are the result of information asymmetries or familiarity. We find no evidence that individual investors possess valuable information. Three separate tests support the idea that simple familiarity drives purchases.

Keywords: Individual Investors, Behavioral Finance, Home Bias

JEL number: G15, F3, D1

*We thank James Guo, Terry Hendershott, Chris Hennessy, Jacob Sagi, Armin Schweinbacher. We also thank members of the finance departments at Berkeley, Duke, and HKUST, the research department at B.G.I., as well as the PBFA Conference. Thuy-Uyen Dam provided some excellent research support. Any mistakes are ours alone. Contact information: Mark S. Seasholes, U.C. Berkeley Haas School of Business, 545 Student Services Bldg., Berkeley CA 94720; Tel: 510-642-3421; Fax: 510-643-1420; email: mss@haas.berkeley.edu.

1 Introduction

Do investors tilt their portfolios towards certain types of stocks? This seemingly simple question has generated, and continues to generate, a large amount of debate among economists. A search for the term "home bias" on EconLit, JSTOR, and SSRN yields over 200 citations. The majority of these papers were written within the last five years.

If investors are tilting their portfolios, what can we learn about portfolio choice? Are there barriers to investment, large frictions, or hidden transaction costs? How well do the choices fit the predictions of a rational expectations model? Are portfolio choices linked to the structure of information and the ability to process it? Or, do individuals gravitate towards stocks with which they are familiar?

This paper undertakes a systematic study of investor holdings in an attempt to answer the questions above. To begin, we consider the multi-asset rational expectations equilibrium model from Brennan and Cao (1997). Their paper considers assets partitioned along national boundaries in an international setting. Without loss of generality, we consider an intra-national setting. Like the earlier work, we consider partitions that are directly related to geography (company location and listing location.) We also consider partitions that are not necessarily related to geography (industry, language group, and cultural group.)

At first glance, our results fit the predictions of the model. Investors who are hypothesized to have better (more) information about a given asset tend to overweight the asset in their portfolios. Assets that are characterized by a wealth of public information do not induce any one group of investors to overweight the asset. Finally, there is weak evidence that an asset's covariance with the market portfolio can predict the level of overweighting.

The initial results lead us to investigate whether individual investors possess better information or simply act as if they have better information (i.e., they have more information or are more familiar with certain stocks.) Investors who hear one company's name repeated frequently may "build-up" priors over time that are consistent with having a precise information matrix. We carry out three tests: i) the portion of an investor's portfolio invested in local stocks does not outperform the portion invested in non-local stocks; ii) events in the distant past are linked to holdings today even though these events do not provide information about today's market; and iii) our finding that investors overweight locally listed stocks has no obvious link to information since all trades take place in cyber-space, on almost identical exchanges, and on exchanges to which investors have equal access.

We exploit the unique market structure in the People's Republic of China (PRC) in order to carry out tests that have previously not been possible. The PRC is divided into thirty-one regions. These regions are culturally different and each has at least one local dialect. Companies and investors are located throughout the country. Our study considers all 945 companies listed in the PRC. Listed companies have a large range of market capitalizations and represent diverse industries. What's more, the PRC has two stock exchanges. Both operate with an electronic limit order book system and have the same proportional trading costs. Company shares trade on one exchange or the other, but are not cross-listed. Access to either market is through a unified and seamless system of computer terminals, which are located in brokerage offices throughout the country.

1.1 Relationship of results to existing research

On the most basic level, our results show that investors tend to overweight locally head-quartered companies in a new sample of data. We call this tendency *pure home bias*. This finding, along with work by Ivkovic and Weisbenner (2004) and Zhu (2003), shows that individuals display the same pure home bias as mutual fund managers and aggregate populations of investors. French and Poterba (1991), Lewis (1995, 1999), and Coval and Moskowitz (1999, 2001) are some well-known papers in this area that look at both international and intra-national holdings. We also confirm that investors do not overweight firms where the amount of public information is high—Ivkovic and Weisbenner (2004) perform a related test with S&P 500 firms.

Our findings are also consistent with an investor having better (more) information about firms that are head-quartered near the investor's cultural/language group. A similar tendency has been documented in Finland among Finnish and Swedish investors. Grinblatt and Keloharju (2001) show that investors are more likely to hold stocks of firms that communicate in the investor's native tongue. We call this tendency *cultural affinity bias* and show it is even more complex than the two-culture system studied in Finland. In fact, cultural affinity bias exists even after dividing investors into thirty-one ethnic/linguistic groups. Although we do find cultural affinity among our investors, we should note that economically and statistically it appears to be one of the smaller biases studied in this paper.

The most surprising results of the paper are the findings related to location of trade. We show that stocks *traded* near where an investor lives receive the highest weights in the investor's portfolio—we call this the *location of trade bias*. While our unique structure allows us to pinpoint location of trade in a manner past studies have not been able to, our results are consistent with existing studies. Papers such as Kang and Stulz (1997), Dalquist and Robertsson (2001), Ahearne, Griever, and Warnock (2004), and Edison and Warnock (2003) all show that U.S. investors tend to hold foreign stocks if they are cross-listed in the U.S. through American Depository Receipts (ADRs). However it is hard to interpret the existing results. One interpretation is that a cross-listing event ameliorates an information asymmetry problem faced by U.S. investors. But the choice of a foreign company to list in the U.S. is not an exogenous event. It is related to the need to raise money, the ability (willingness) to meet accounting standards, and global marketing goals. Studies such as Bekaert and Harvey (2000) and Bekaert, Harvey, and Lumsdaine (2002) define market liberalization events by using cross-listing dates. Our study offers a clean environment that controls for these effects. A Beijing-headquartered company that is listed in Shanghai is subject to the same accounting standards as a Beijing company that is listed in Guangdong (on the Shenzhen exchange). Investors and companies have a common language and operate in the same time zone. There is no cross-listing in the PRC, nor may companies switch exchanges.

The finding that location of trade is the biggest factor in an individual's portfolio leads to two other contributions of this paper. The first contribution revolves around a series of tests designed to test if portfolio choices are the result of information or familiarity. The fact that investors in our sample have unfettered access to very similar electronic exchanges makes it difficult to connect location of trade bias to information asymmetries. More importantly, our findings shed light on recent evidence that location of trade affects asset prices. We discuss this in section 1.2 directly below.

1.2 Location of trade, asset pricing, and clientele effects

A major motivation behind our study is recent evidence that location of trade affects asset prices. Froot and Dabora (1999) study "Siamese twin" companies. Such companies have two separate shares that trade, they pool their cash flows, and their charters dictate current and future divisions of cash. Thus, with integrated markets, the twin stocks should move together. The authors find that this is not the case. For example, Royal Dutch Petroleum trades primarily in the U.S. and the Netherlands. Shell Transport and Trading trades primarily in the U.K. The authors note that "when the U.S. market moves up relative to the U.K. market, the price of Royal Dutch (which trades relatively more in New York) tends to rise relative to the price of its twin Shell (which trades relatively more in London)." The authors reject that voting rights, currency fluctuations, ex-dividend timing, and tax differences can explain all of their findings. In a more recent paper, Chan, Hameed, and Lau (2003) perform an event study surrounding the de-listing of Jardine companies from the Hong Kong Stock Exchange (and the subsequent migration of trading volume to the Singapore Stock Exchange.) Before the de-listing, Jardine companies had a high covariance with the Hong Kong index and a low covariance with the Singapore index. After the de-listing, the same companies had a low covariance with Hong Kong and a high covariance with Singapore. The authors cannot explain their results by looking at possible relocation of core businesses, currency movements, tax distortions, or time-varying covariance.[1]

Both studies—Froot and Dabora (1999) and Chan, Hameed, and Lau (2003)—suggest that country-specific investor sentiment can explain their results. However, neither group offers evidence beyond ruling out other explanations. One can interpret these recent findings as investors being partially segmented—see Chan, Hameed, and Lau (2003, p. 1222). In other words, there is a *clientele* effect. Different clients operate in different locations and affect stock prices/betas differentially. Arbitrageurs appear limited in their ability to equalize prices across markets, as in Shleifer and Vishney (1997). If clientele effects exist, we should be able to measure differences at the stock-holding level. Neither of the papers mentioned above attempt to do this due to the difficulty of obtaining shareholder records and demographic information about each shareholder. This paper attacks the problem directly by using account-level data.[2]

The paper proceeds as follows. In Section 2 we review the testable implication of the rational expectations model. Section 3 describes the data and market structures that make this study possible. Section 4 outlines our methodology and presents the empirical results. Section 5 presents three series of tests designed to determine if information asymmetries or familiarity is driving our results. Section 6 tests alternative economic models and performs robustness checks. A brief conclusion is provided in Section 7.

[1] In addition, Bedi, Richards, and Tennant (2003) expand the research of dual-listed companies. They document twelve new structures, show that price divergence has remained pervasive, show that betas change depending on listing, and conduct an event study of companies with unified share structures.

[2] To be fair, our paper looks at clientele holdings directly as we do not have a natural experiment involving cross-listing. For clientele effects to explain the results of Froot and Dabora (1999) and Chan, Hameed, and Lau (2003) investors would need to be segmented by location of trade (or self-segment by location of trade.) Of course, a sufficient condition would also include frictions that keep rational arbitrageurs from taking unlimited positions in order to drive prices to convergence.)

2 Multi-asset rational expectations model

In order to understand portfolio choices, we test predictions from the Brennan and Cao (1997) multi-asset noisy rational expectations model. Their paper considers M risky assets partitioned along national boundaries in an international setting. Each asset, indexed $m = \{1, \ldots, M\}$, represents a different country's equity market index. A continuum of investors is also partitioned into M groups based on the same national boundaries.

The model has T trading sessions held at times: $t = \frac{0}{T}, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}$. Asset payoffs are random, take place at time 1, and there is no intermediate consumption. Before each trading session, investors receive a $M \times 1$ vector of private (non-public) signals about the asset payoffs and a $M \times 1$ vector of public signals about the asset payoffs. While this dynamic models allows one to test implications regarding cross-border flows and asset prices, Feng and Seasholes (2004) already use a simplified version to test for correlated trading by location. We therefore concentrate on the implications regarding portfolio holdings.

Without loss of generality, we consider M risky assets in an intra-national setting. Like earlier work, we consider partitions that are directly related to geography (company location and listing location.) We also consider partitions that are not necessarily related to geography (industry, language group, and cultural group.) We highlight two key assumptions from the original paper.[3]

> **Assumption A, cumulative information advantage:** A partition of investors, m, has better (more) information than a partition of investors k (where $k \neq m$) about the payoffs to asset m. The partition of investors m has worse information than k about the payoff to asset k. Furthermore, the marginal informational (dis)advantage from a private signal in time period $\frac{t}{T}$ is small relative to the cumulative information advantage from all private signals received over the time period $[\frac{0}{T}, \frac{t}{T}]$.

[3]For the exact specifications of these assumption, please see p.1861 of Brennan and Cao (1997). There is also a necessary and sufficient condition that the conditional covariance of the asset m's payoff with the market portfolio be positive (see p.1863-1863.)

Assumption B, symmetric information endowments: Assume there is no asymmetry in partition m's precision matrix regarding asset payoffs. One way to obtain symmetric information endowments is to assume investors in partition m receive private signals only about asset m and they receive no private signals about asset k (where $k \neq m$.)

The model produces the following testable predictions:

H_A **(home bias):** Under Assumptions A and B, a partition of investors should hold more (less) than their share of an asset for which they have a cumulative information (dis)advantage.

H_B **(public information):** For an asset m, where public information outweighs private information, we do not expect to find any partition of investors holding more than their share of the asset.

H_C **(unrelated investor partitions):** For an asset m and a partition of investors $n = \{1, \ldots, N\}$ that is unrelated to private information, we do not expect to find any partition holding more than their share of the asset.

H_D **(covariance with market portfolio):** Under Assumptions A and B and for a given level of information asymmetry, home bias will be greater for assets that have a higher covariance with the market portfolio.

Since an investor's information set cannot be measured by the econometrician, the testable hypotheses must be related to observable characteristics of firms and individuals. Table 1 outlines eleven testable hypotheses related to H_A and three related to H_B. We do not list anything regarding H_C since there are infinite unrelated partitions. Finally, H_D is suggested by Brennan and Cao (1997), thought it is not tested in their paper. The next section describes the data we use to test these hypotheses.

3 Data and market structure

We use account-level data to investigate investor portfolio composition. Our data come from individual brokerage accounts in the People's Republic of China (PRC). By and large, we focus on holdings at one point in time, 01-June-2000, although other time periods between 1999 and 2000 are used as controls.

3.1 Cultural divisions in the PRC

The PRC offers a unique opportunity for social scientists to study geographic and cultural differences. The country is divided into thirty-one regions (provinces, autonomous regions, and municipalities); Figure 1 shows the location of the regions. Officially there are fifty-six ethnic groups in the PRC. The largest ethnic group is the Han, who comprise over 90% of the population. The Manchu, Mongolian, Tibetan, and Uygur are some other well known groups. All regions speak Mandarin (called "Putonghua" or "common language.") Regions have at least one local dialect. The best known of these dialects is Cantonese (or "Guangdonghua") which is spoken in Guangdong province, Hong Kong, and many overseas Chinese communities.

3.2 Equity markets in the PRC

The structure of equity markets in the PRC offers an unparalleled opportunity for studying investor portfolios. We have a clean, laboratory-like research design. The country uses an electronic, open limit order system. Investors can enter their own trades through computer terminals that show the current queue five deep on both the bid and ask side. The market has uniform and low transaction costs. One of the strengths of this study is that we are assured that investors do not have other brokerage accounts. Hong Kong, Macao, and Taiwan are not included in this study for political, legal, and regulatory regions. The typical PRC investor cannot easily invest outside the PRC. Since mutual funds are relatively new to the PRC, we effectively know investors' entire equity portfolios.

Exchanges: There are two stock exchanges in the PRC. One is in Shanghai and the other in Guangdong (in the city of Shenzhen). The Shanghai Stock Exchange uses six-digit tickers, while Guangdong (Shenzhen Stock Exchange) uses four-digit tickers. Investors have equal and unfettered access to either exchange.

Listed companies: We identify 945 listed companies with traded common shares on 01-June-2000. Of these, 485 companies are listed in Shanghai and 460 in Guangdong. At least one listed company has its headquarters in each of the thirty-one regions. Given thirty-one regions and two exchanges, we can classify firms and holdings into one of sixty-two region/exchange bins. Although Table 1 lists a number of testable hypotheses, we do not have the following firm information at this time: language spoken by CEO, distribution of firm products, leverage, and fraction of intangible assets.

IPO data: Initial Public Offering (IPO) data contain the issue date, name of the bank that managed the IPO, and location of the bank. This data is proprietary and collected from a private source in the PRC. Data are available for 778 of the 945 companies in our sample.

Brokerage accounts: A brokerage firm (the firm) has branch offices (branches) throughout the country, region, or city. Many brokerage firms are regionally focused. Individuals open accounts at a branch office and then place all of their trades *through this one branch.* Thus, there is a critical difference in our study between brokerage firms (our data are from one firm) and branch offices (our data come from fifteen different branches.) For a full description of brokerage accounts in the PRC, please see Feng and Seasholes (2003, 2004).

Our data contain each investor's internal passport number (or "NIC" number). This number can be decoded and gives the individual's birth date and gender, as well as where the individual is registered to live. Given the fact that investors place all trades through one branch, we assume investors currently live in the region where the branch is located (we call this the "home region"). Our investors live in seven of the thirty-one regions. In cases where the registration location is not the same as the current region, we assume that the investor has joined the vast internal migration in the PRC. We assume the NIC number gives us the region where the investor was born (or "birth region"). The effects of these assumptions are discussed further in Section 6. Our data contain at least one investor who was born in each of the thirty-one regions. Given the seven home regions, individuals can be classified in one of 217 birth-region/home-region bins.

3.3 Overview statistics

Table II presents an overview of the data, which is collected from fifteen branch offices located in seven regions (Column 1 and Column 2.) The two regions with stock exchanges (Guangdong and Shanghai) both have multiple branch offices.

Overall, we have the holdings of 51,218 individuals on 01-June-2000 (Table II, Column 3). We focus on a single date in the middle of the sample period to study portfolio holdings and save other dates for robustness checks. This focus is conservative since it lowers potential statistical power. On the other hand, using one date simplifies statistical tests, because we do not have to worry about observations being correlated across time. The median investor holds only three stocks (Column 4.) The average investor holds 3.47 stocks (not reported). This gives us a total of 177,783 investor/stock positions. In Columns 5, we see the average portfolio value is RMB 136,777. This average value is approximately equal to USD 17,097 if we use a rough exchange rate of RMB 8:USD 1. Notice that richer regions such as Guangdong and Shanghai have higher average portfolio values than poorer regions such as Heilongjiang. The distribution of portfolio values is skewed, as we can see by the median values in Column 6. The total value of all holdings used in this study is given in Column 7 and is over RMB 7 bn (or slightly less than USD 1 bn.) Appendix 1 provides some overview statistics such as population, GDP per capita, and household income for each of the thirty-one regions.

4 Methodology and results

4.1 Reference portfolio and unit of analysis

Statistical tests of portfolio tilting entail a joint hypothesis that: i) investor portfolio weights are equal to the weights in a reference portfolio; and ii) we choose the *correct* reference portfolio. Choosing a reference portfolio is difficult, fraught with uncertainty, and controversial. Dahlquist and Robertsson (2001) provide a nice discussion on this point. A logical reference portfolio is the CAPM portfolio as described in Sharpe (1964) and Lintner (1965). Alternative portfolios are considered later in this paper.

We call ω_i^m the weight of stock i in investor m's portfolio.[4] We call ω_i^* the weight of the stock(s) in the market. The difference is between these two weights is our measure of the investor's over/underweighting of the stock(s):

$$\text{over/underweighting} \equiv \omega_i^m - \omega_i^* \tag{1}$$

[4]Note i can also represent a group of stocks and m can also represent a partition of investors.

4.2 Preliminary results: partitions with an information advantage

We begin by looking at holdings of stocks where some investors may have a information advantage. From the list of testable hypotheses in Table I, we choose three (A.1, A.2, and A.3) that are all related to geography and for which we have the necessary data. We use a binary methodology to classify holdings as coming from (or not coming from) the region where an investor lives. This methodology is similar to most international studies of home bias that group holdings as coming from (or not coming from) an investor's home country: Coval and Moskowitz (2001) code holdings as within (or outside of) a 100-kilometer radius around the investor; Ivkovic and Weisbenner (2003) use a 250-mile radius; and Grinblatt and Keloharju (2001) classify holdings by municipality in Finland.

Table III, Panel A shows that investors tend to overweight companies that are headquartered in the region where the investor currently lives. On average, 19.53% of investor' portfolios come from the region where they currently live (see the average value at the bottom of the second column.) On average, only 11.28% of the market is headquartered in the region where the investor lives. The difference between these two numbers is 8.25%, and the positive value indicates overweighting. We see that the average overweighting due to pure home bias is positive when investors are grouped by the seven regions where they currently live. It is also positive when investors are grouped by the fifteen branches offices that provided our data (not shown). At this point we only present overview statistics and save tests of statistical significance for the next section. This said, we note that with 51,218 investors, 8.25% is statistically significant at all conventional levels.

Table III, Panel B shows the cultural affinity bias. On average, an investor holds 16.92% of his or her portfolio in companies with headquarters in the region where the investor was born. On average, 9.33% of the market capitalization is from birth regions. Therefore, the average overweighting due to cultural affinity bias is 7.58%.

To test hypothesis A.3 from Table I we consider the portfolios of investors who live in Guangdong and Shanghai since this is where the two stock exchanges are located. In Table III, Panel C we see about 51.57% of the total PRC market capitalization is traded on the Shenzhen Stock Exchange in Guangdong Province. The remaining 48.43% is traded in Shanghai. However, investors from these regions have over 80% of their holdings in locally-listed stocks. This makes an average of 34.28% overweighting due to the location of trade bias.

The overweighting due to the location of trade bias is huge and very, very surprising. This is the first time (that we know of) that financial economists have been able to estimate location of trade effects related to actual stock holdings. Our surprise stems from the fact that both exchanges are electronic. Investors have equal access to either exchange. There is no difference when buying shares on one exchange or the other. In fact, the 34.28% is so large that it makes the prudent reader skeptical. Maybe 34.28% is really just another manifestation of pure home bias? As we can see in Table III, pure home bias is especially pronounced in Shanghai and Guangdong. Maybe the location of trade findings are related to a cultural affinity bias? To check these possibilities, we turn to a multivariate regression framework in Section 4.5 below.

4.3 Preliminary results: partitions without an information advantage

Our next test focuses on stocks where an information advantage is hypothesized not to exist. From Table 1, we chose hypotheses B.1 and note that these stocks should not be disproportionately held by any one group of investors. To test this hypothesis we choose the largest 237 stocks ranked by market capitalization (this is the top quartile of all listed companies.) We then consider three different partitions of investors: i) all investors in our sample; ii) investors divided by the seven provinces where they currently live; and iii) investors divided by the fifteen branch offices in our sample. Table IV shows our results. In Panel A, we see that the largest 237 stocks account for 57.94% of the total market capitalization. Investors in our sample allocate an average of 58.13% of their portfolios to these stocks. In Panel B, we partition investors into the seven province where they currently live. On average, investors allocate between 53.69% to 62.59% to the largest 237 stocks. In Panel C, we partition investors into the fifteen branch offices in our sample and show overview statistics. An interesting trend is worth noting. As we separate into more and more partitions, the probability of any one group holding close to the sample average 58.13% of its portfolio in the top 237 stocks drops. This should not surprise us since we know that individuals hold about three stocks in their portfolios. If we were to partition investors into the 51,218 individual accounts, some would have 0% of their portfolios allocated to the largest 237 stocks while other would have 100% of their portfolios allocated to the same stocks.

4.4 Alternative measures of overweighting

We have chosen to use differences between observed weights and a reference portfolio as shown in equation (1). We could also have looked at the ratio of observed weights and the reference portfolio by using:

$$\text{over/underweighting ratio} = \frac{\omega_i^m}{\omega_i^*} - 1 \tag{2}$$

From Table III, we see this new ratio is also positive in all cases. The interpretation is a bit different. In all three biases (pure home bias, cultural affinity bias, and location of trade bias) investors hold approximate twice as much (100% more) than the reference portfolio. While the location of trade bias no longer dominates using the ratio measure, its value is capped from above at one (if 50% is listed on one exchange, investors can only add another 50%.) Additionally, putting 3.15x the reference weight in a small region like in Hubei is interesting, but tells us little of wealth committed by investors. For these two reasons, we stick with our difference measure in equation (1) and note that statistical significance remains with equation (2) albeit with a different interpretation.

4.5 Regression analysis

We turn to a regression analysis with two main goals. First, we want to double-check the results from Table III—especially the surprising results related to the location of trade bias. Second, we want to determine whether the three documented biases are, in fact, independent. It is entirely possible that the pure home bias, cultural affinity bias, and location of trade bias are simply manifestations of a common bias or common factor.

For each of the 51,218 investors (m), we calculate the percentage of the portfolio that is held in stocks from each of the sixty-two region/exchange combinations and call this measure $\left(\omega_{r,e}^m\right)$. We also calculate the percentage of total market capitalization that is listed in each of these sixty-two region/exchange combinations and call this $\left(\omega_{r,e}^*\right)$. We then calculate the over- or underweighting by taking the difference of the two values.

Our procedure allows us to look at pure home bias, cultural affinity bias, and location of trade bias together. In total, there are 3,175,516 over/underweighting observations.[5] Clearly,

[5] Calculated as: $3,175,516 = 51,218\ \textit{investors} \times 31\ \textit{regions} \times 2\ \textit{exchanges}$.

we do not really have this many data points since most investors only hold three stocks at a time. Therefore, we correct our standard errors to reflect the true sample size of 149,691 data points.[6]

We regress our over- or underweighting measure $\left(\omega_{r,e}^{m} - \omega_{r,e}^{*}\right)$ on one of three indicator (dummy) variables. The dummy variables effectively serve to partition investors and stocks into the $M = 62$ groups based on region and exchange.[7] The first indicator $\left(Dum_{r,e}^{Home=HQ}\right)$ equals one if the region where is the stock is headquartered is the same as the region where the investor currently lives (and equals zero otherwise). The second indicator $\left(Dum_{r,e}^{Born=HQ}\right)$ equals one if the region where the stock is headquartered is the same as the region where the investors was born. The third indicator $\left(Dum_{r,e}^{Home=Listing}\right)$ equals one if the listing exchange of the holding is located in the same as the region where the investors currently lives.

$$\omega_{r,e}^{m} - \omega_{r,e}^{*} = \gamma_1 Dum_{r,e}^{Home=HQ} + \gamma_2 Dum_{r,e}^{Born=HQ} + \gamma_3 Dum_{r,e}^{Home=Listing} + \varepsilon_{r,e}^{m} \quad (3)$$

We do not include a constant since we fully realize there is an adding up constraint as far as an individual's portfolio weights are concerned. For example, an investor who holds one Beijing-headquartered stock that is listed in Shanghai has 100% of his or her portfolio in the Beijing/Shanghai exchange bin. Since companies in this bin only represent 3.35% of total market capitalization, we can say that the investor has overweighted the bin by 96.65%. At the same time, the investor has underweighted *all other* region/exchange combinations. In fact, the investor has an average 1.58% underweight in all other region/exchange bins.[8] We do not want to measure the difference between the overweighting and the average underweighting. We simply want to measure the difference between the investor's portfolio and the reference portfolio.

Table V presents the results of four regressions in a manner than can easily be compared with Table III. The first three regressions repeat the overview results from Table III and provide a measure of statistical significance. The most interesting point in Table V is regression 4. Again, we see the surprising result that the location of trade bias is dominant and economically significant. Investors tend to overweight their portfolios by 30.36% due to the location of trade bias. Pure home bias overweighting is 6.49% and remains statistically

[6] We have 149,691 total investor/region/exchange holding combinations in our dataset. Using this number is conservative since we group multiple holding positions by same investor into the same investor/region/exchange bin.

[7] There is a slight abuse of notation here as m may stand for one investor or one of the 62 region/exchange partitions.

[8] We get 1.58% by dividing the 96.65% underweighting by the 61 remaining region/exchange combinations. The 96.65% underweighting comes from the fact that the sum of all over- and underweightings is zero (by definition) for a given investor.

significant. Economically, this number is only a fifth as large as the location of trade bias. The cultural affinity bias is economically and statistically much smaller than either of the other two biases.

Regression fit: How well does regressing holdings on dummies fit our data? If we consider all 3,175,516 investor/region/exchange combinations, the fit is not good since investors only have holdings in 149,691 of the bins. However if we only consider investor/region/exchange combinations with positive investor holdings, the R^2 ranges from 0.0843 for regression 2 to 0.1871 for regression 4. In other words, we get a fairly good fit by simply by guessing that an investor spreads his/her stocks evenly over local stocks.

Clearly, using a censored sample of 149,691 non-zero bins changes the regression coefficients presented in Table V. Although not reported, the new coefficients follow a similar pattern to the one shown in Table V. In particular, when estimating equation (3), the coefficient on the location of trade dummy is the largest and the coefficient on the cultural affinity dummy is the smallest.[9]

4.6 ANOVA analysis

Since Table III and Table V provide surprising results concerning the location of trade bias, we retest our results with ANOVA analysis. We run a regression that is essentially the same as equation (3). The only difference is that an overall constant is included this time so that the sum of squares of the first right-hand side variable does not pick up the variance associated with the investor's adding-up constraint (see Section 4.5 above for a brief discussion of the adding-up constraint.)

Table VI shows that the indicator variable associated with the location of trade bias explains over three times as much of the variance as does the indicator variable associated with pure home bias $\left(\frac{100.58}{30.12} > 3\right)$. It explains more than forty times as much of the variance as the cultural affinity bias $\left(\frac{100.58}{2.38} > 40\right)$. Results are significant at all conventional levels.

[9] We also run the regression using the ratio form of our over/underweighting measure from equation (2). A similar high-level of statistical significance is found. Other robustness checks are discussed in Section 6.

4.7 Multifactor model

The regression and ANOVA analyses provide consistent and statistically significant results. The results are so surprising that we run one more test in an effort to make sure we are interpreting our results correctly. We regress the returns from investor portfolios on the overall market index, regional indices, and listing/exchange-based indices. The market index is a weighted average of all 945 companies in our sample. The regional indices are weighted averages of stocks that are headquartered in each region. The two listing/exchange-based indices are weighted averages of the stocks listed on the two exchanges. Our goal is to see if investors choose portfolios that expose them to location-based factors.

$$r_t^m = \alpha + \beta \cdot r_t^{mkt} + \Gamma_1 \cdot r_t^{home} + \Gamma_2 \cdot r_t^{birth} + \Gamma_3 \cdot r_t^{exch} + \varepsilon_t^m \tag{4}$$

To generate the investor portfolio returns, we calculate the weight of each stock in each investor's portfolio (regardless of how many different stocks an investor might have). We then calculate the weekly return to the investor's portfolio over the *following* twenty-six weeks. We re-balance portfolios every six months for one year. Thus, we have a panel of two six-month periods (with twenty-six weeks per period) times the number of investors (or groups of investors.) To control for clustering, we form portfolios of the investors' portfolios: all investors who live in one region, all investors who were born in one region, or all investors who live near one of the two stock exchanges. Our results do not change when we consider individual investor portfolios separately and control for clustering of the residuals.

Table VII presents the results from the multifactor model. In regression 1 we regress a portfolio of all investor returns on a constant and the returns of the market. Regression 1 provides some comforting results. On average, investor portfolios have a beta of one and an alpha near zero.[10]

We then regress investor portfolio returns (portfolios of investors based on where they currently live) on a constant, the returns of the market, and the associated regional index (again based on where the investors currently live). Table VII, regression 2 shows that investor portfolios load positively (and significantly) on the home-region portfolio. This is

[10]Regression 1 uses fifty-two data points (calculated as fifty-two weeks of data—one year—times one group of investors.) The R^2 of the regression is 0.9724—high since investors, as a group, basically track the market. We consider using only fifty-two data points a conservative test. Alternatively, we could have used 2,663,336 data points (51,218 investors times 52 weeks) and controlled for contemporaneous correlation of observations across investors.

our fourth piece of evidence of pure home bias. In a similar manner, Table VII, regression 3 shows that investor portfolios load positively (and significantly) on the birth-region portfolio. Similarly, this is our fourth piece of evidence of cultural affinity bias.

Finally, we perform regressions for investors who do not live in a region with an exchange (regression 4a) and those who do (regression 4b). We see those who do not live in a region with an exchange again have market betas close to one. However, the portfolio returns of those who live in Guangdong and Shanghai load heavily on their local exchange index. The 0.4768 coefficient shown in Table V, regression 4b *is after controlling* for the covariance of an investor's portfolio with overall market movements.

4.8 Return covariance and home bias

The final testable hypothesis from the rational expectations model (H_D) stats that for a given level of information asymmetry, pure home bias will be greater for assets with higher covariance with the market portfolio. We would like to test this directly, but our sample contains investors who currently live in only seven regions. Therefore, we consider the investors' birthplaces to give us a sample size of thirty-one. The small sample size and the knowledge that cultural affinity bias is the smallest we study here, makes this test weak.

For each of the thirty-one regions in the PRC, we calculate the covariance of the regional index with the overall market. The regional index is the weighted average return of companies from the region. We also calculate the cultural affinity overweighting for each region. We then regress the overweighting measure on the covariance and find the regression coefficient is positive as the theory predicts. However, the t-stat is not significant at conventional levels.

4.9 Economic costs

After documenting the biases, it is natural to ask how costly they are to an investor. We could calculate welfare loss in a manner similar to Brennan and Torous (1999). However, this requires specifying investors' utility functions and risk aversion parameters. Another method would involve testing the efficiency of various portfolios. We could measure how "far" a portfolio is from the mean-variance frontier. However, methods of this sort often require us to specify expected returns.

We realize that tests of economic costs can be controversial, and opt for a straightforward comparison of Sharpe ratios. The advantage lies in the simplicity. We use market weights as of 01-June-2000, form a market portfolio, and calculate the weekly return for the next twenty-six weeks. We divide the average weekly return over this period by the weekly standard deviation. We repeat the procedure for portfolios formed from stocks on each exchange and from each region. Our results are shown in Table VIII.

Table VIII, Column 2 shows a weekly Sharpe ratio of 0.1356 for the entire market. The average value is 0.1253 for the exchange indices. This value is 0.0103 less than the market Sharpe ratio. In Column 3 we interpret this difference as a "cost" to a mean-variance investor. We also calculate the average Sharpe ratio for a value-weighted portfolio of five random stocks.

It is clear from Table VIII that for small investors who hold few stocks, the lack of diversification is a major cost. For investors or fund managers who hold many stocks, the location of trade bias also appears costly. To see this, consider that the typical portfolio has 30.26% more weight on the local exchange than the reference portfolio (Column 4). If we multiply this number by the cost in Column 3, we get 31 bp (basis points). For pure home bias, a similar calculation yields 20 bp. We have expressed cost in units of *price of risk*. On this basis, our rough estimation is that location of trade bias is 50% more costly than pure home bias (31 bp vs. 20 bp.)

In this section, we show that location of trade appears to be the dominant factor affecting individual portfolio choice. Our tests take place in a unified market system with very low transaction costs. Despite open access to either stock exchange, we show that some investors tend to self-segregate. Section 5 looks at how the information environment may or may not affect investors. Section 6 then tests other hypotheses that might explain the observed segmentation.

5 Information or familiarity

Section 4 shows that investors tilt their portfolios towards certain stocks. Investors either have better information about these stocks or they are simply more familiar with the stocks. This section is devoted to testing which effect is driving the portfolio choices of individual investors.

5.1 Performance of the local-portion of a portfolio

Our first test is to check the returns each investor makes on the local portion of his portfolio compared to the non-local portion. Coval and Moskowitz (2001) and Ivkovic and Weisbenner (2004) show that the local part of U.S. investors' portfolios tend to outperform the distant part of their portfolios. Zhu (2003) does not find this result when looking at the same data as Ivkovic and Weisbenner (2004).

For comparability, we closely follow the methodology as Ivkovic and Weisbenner (2004). We regress investors' raw portfolio returns, $R^i_{t,t+k}$, over a k-week horizon ($k = 26$) on the portfolio weights of local and non-local investments (no constant). We calculate local investment three different ways: i) the percent of the portfolio located where the investor currently lives (pure home bias); ii) the percent of the portfolio located where the investor was born (cultural affinity bias); and iii) the percent of the portfolio listed on the exchange in the region where the investor currently lives (location of trade bias).

Table IX presents our results. In panel A we see the sign of the locality difference is negative. This result is consistent with investors not having valuable information about local stocks, yet choosing to hold them. The sign of the locality difference is negative in Panel B where local is defined as a company from the same region where the investor was born. Interestingly, the sign of the locality difference becomes positive in Panel C when we consider stocks listed near where an investor currently lives. Since the location of trade bias is so dominant, we might interpret this result and local information.

When analyzing statistical significance, we realize a problem exists. It is unlikely that we actually have 51,218 independent observations. After all, there are only 945 possible stocks to choose from and many investors hold just a few stocks. We propose allowing independence across groups of investors but not necessarily within a given group (clustering). Of course, it is not clear how to group investors. One of the least restrictive groupings allows for clustering by the dominant stock in an investor's portfolio. A more restrictive assumption is that all investors from the same region or same birth-region/home-region are non-independent observations.

We note a disturbing trend in Table IX. The standard errors increase by a factor of 2.5 to five times when we allow clustering. This is worrisome for the following reasons. Our sample of data is comparable to the Ivkovic and Weisbenner (2004) sample. We have 51,218 accounts; they use 27,032 households. We have a sample of 945 stocks; their sample is likely to have on the order of 5,000 stocks. One of their main results is that "the average household generates an additional annualized return of 3.2% from its local holdings relative to its non-local holdings." The t-stat is 4.57 as reported in Table 5 of their paper. The t-stats in our Table IX are about the same magnitude as theirs. Using robust (White) standard errors does not change significance much which is not surprising. The worry in this type of regression is not generalized heteroskedasticity but rather correlated observations. As we can see in our Table IX, any allowance for clustering makes the results insignificant.

Conclusions from portfolio performance tests: The conclusion from this section is clear. Investors are no better at picking local stocks than they are at picking non-local stocks. Yet, they invest much of their wealth in stocks with which they are (presumably) more familiar.

5.2 Long-lived familiarity effects

Our second test identifies a situation where investors may have become familiar with a stock in the past. But, it is hard to argue that the situation gives investors relevant information about the current investment environment. The test involves looking at the location of the investment bank that ran the initial public offering (IPO.) We are interested in testing whether investors who were more exposed to information about the company during the IPO process are more likely to hold the company's stock today. On average, stocks in our sample have been listed for four to five years and these stocks have a turnover of approximately two times per year. We hypothesize that the IPO process generates information about soon-to-be-listed companies only. Formally, the null hypothesis is that events in the past (e.g., four years ago) should have no effect on holdings today.

We obtain a proprietary list of initial public offering dates and the names of the investment banks who managed the IPO. We know the location of each bank. The data cover 778 of the 945 companies in our sample and provide the location of the investment banks that managed the IPO.

Pure home bias: For each of the 778 companies, we count the total number of shares held by all 51,218 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 company's headquarters. We call this our aggregate measure of pure home bias ($\Omega_i^{\text{pure home bias}}$) for stock i. It is important to note that we do not have the complete holdings of each listed-firm. Therefore these results in this section contain sampling noise.[11] For the pure home bias test, we consider only the 388 firms where we have at least one investor in our sample, who lives in the same region as the firm's headquarters, and holds some of the firm's stock. One can think of our test as follows: we take home bias as given and then test for differences (if any) in a home bias measure across firm characteristics. For each of the 388 firms, we regress our aggregate measure of pure home bias on a constant and an indicator variable ($Dum_i^{\text{I-bank}}$). The indicator variable ($Dum_i^{\text{I-bank}}$) equals one if the investment bank the managed firm i's original IPO is headquartered in the same region as firm i is headquartered.

$$\Omega_i^{\text{pure home bias}} = \alpha + \gamma \cdot Dum_i^{\text{I-bank}} + \varepsilon_i \tag{5}$$

The results are presented in Table X, Panel A and are quite stunning. The fraction of shares held in home region-firms almost doubles if the investment bank was headquartered in the same region. To see this doubling we look at the regression estimates. Holdings are 0.2652 in firms where the investment bank was not local and are 0.4545 (=0.2652+0.1893) in firms where the investment bank was local.

Cultural affinity bias: Table X, Panel B repeats a similar test for cultural affinity bias. In this test, we count the fraction of shares held by investors who were born in the same region as the company's headquarters and call the measure $\Omega_i^{\text{cultural affinity bias}}$. Due to the limited nature of our sample, we consider the 595 of 788 firms with a positive measure of $\Omega_i^{\text{cultural affinity bias}}$. We again run a regression on a constant and our indicator variable ($Dum_i^{\text{I-bank}}$). The fraction of shares in the birth region is more than double the average fraction after considering the presence of the investment bank.

$$\Omega_i^{\text{cultural affinity bias}} = \alpha + \gamma \cdot Dum_i^{\text{I-bank}} + \varepsilon_i \tag{6}$$

Location of trade bias: Finally, Table X, Panel C tests long-lived effects and location of trade bias. We now calculate the fraction of shares in our sample that are held by investors who live in the same province (Shanghai or Guangdong) as the firm is listed. We call the measure $\Omega_i^{\text{location of trade bias}}$. Again we consider only the 667 of the 778 firms for which this

[11]Our sample is representative of the overall market in many dimensions—see Feng and Seasholes (2003). However, we know that investors currently live in only seven of the thirty-one regions.

measure is positive. We regress our measure on a constant and an indicator variable. This time, the indicator variable ($Dum_i^{\text{I-bank}}$) equals one if the investment bank is headquartered in the same region as the stock is listed. We see the presence of a local investment bank almost triples the fractions of shares held (0.1689 compared with 0.1689+0.2735=0.4424).

$$\Omega_i^{\text{location of trade bias}} = \alpha + \gamma \cdot Dum_i^{\text{I-bank}} + \varepsilon_i \tag{7}$$

Conclusions from tests of long-lived effects: We see that living near the investment bank that originally managed a company's IPO, greatly increases one's chances of holding the stock. We hypothesize that living near such an investment bank once helped investors to become familiar with the company. We argue that any information at the time of the IPO is not relevant in today's investment environment. The results in Table X indicate that familiarity, not information asymmetry, is responsible for portfolio holdings.

5.3 Field study in the PRC

We visited brokerage offices in a number of cities in the PRC, watched trading behavior, interviewed brokerage firms, and collected newspapers. The goal of the field study was to asses the cost of gathering stock information in the PRC. While any study of information environments in the real world is bound to be selective, we make the following observations.

Brokerage offices: Brokerage offices in the PRC have large, electronic boards that update stock prices throughout the day. These boards list stocks from both exchanges regardless of where the office is located. For example, brokerage offices in Shanghai show prices of *both* Shanghai-listed and Guangdong-listed stocks.

Trading terminals: Investors typically trade through terminals located in brokerage houses. These terminals offer equal access to the stocks listed in Shanghai and Guangdong. The availability of information (past stock prices, volume, order queue, etc.) is the same for stocks listed on both exchanges.

Business newspapers: There are three main business newspapers in the PRC: China Securities Journal, Securities Times, and Shanghai Securities News. The offices of the three papers are located in Beijing, Guangdong(Shenzhen), and Shanghai. All three papers provide stock information on both Shanghai-listed and Guangdong-listed stocks.

Local newspapers: Local newspapers focus on local companies and local events. This might explain pure home bias but not location of trade bias. Local newspapers that list stock prices tend to list prices from both exchanges.

The PRC is a country that uses electronic exchanges and investors have equal access to either exchange. It is hard to argue that there are large informational asymmetries driving the portfolio weights we measure—especially the location of trade bias.

6 Tests of alternative models and robustness checks

We now turn to existing models that attempt to explain portfolio biases. We test whether the predictions of these models can explain the dominance of the location of trade bias that we measure. We also retest our earlier results for robustness.

6.1 Transaction costs

Direct and high transaction costs: Stulz (1981) presents a model in which it is costly for investors to hold foreign securities. Our research design has specifically controlled for costs by examining an intra-national setting with uniform transaction costs. Given the high turnover of international portfolios, Tesar and Werner (1995) posit that variable transaction costs are an unlikely explanation for home bias. The PRC also has high turnover, which leads us to believe that direct transaction costs cannot explain the portfolio tilting shown in this paper.

Indirect and small transaction costs: It is possible that investors experience very small transaction costs when switching between trading Shanghai-listed stocks and Guangdong-listed stocks. The costs may be related to execution time. During our sample period, brokerage offices maintained two accounts per individual (one account for holding stocks from each exchange). This system was much like the American system of having a checking and savings account at the same bank. Maybe investors didn't like switching back and forth between accounts.

If small transaction costs are driving our results, we can formulate two hypotheses about the investors who do not live near a stock exchange (i.e., not in Guangdong nor in Shanghai). In aggregate, these "non-exchange" investors should hold roughly similar amounts of Shanghai-listed and Guangdong-listed stocks. At the individual level, the majority of investors should show a distinct preference for trading on one exchange over the other.

We select the 30,171 investors in our sample who live in the five non-exchange regions (Beijing, Heilongjiang, Hubei, Shandong, and Sichuan.) In aggregate, 56% of the stocks held by these investors are listed in Shanghai and 44% are listed in Guangdong.

At the individual level, 16,564 of the 30,171 investors hold 100% of their portfolio in stocks from one of the two exchanges. This high fraction (16,564/30,171) is misleading. 8,643 of the investors hold only one stock and therefore we can say nothing about a preference for one exchange over the other. Another 4,271 of the 16,564 hold only two stocks. Thus, there is approximately a 50% chance that both stocks will be listed on the same exchange. And 1,935 of the 16,564 hold only three stocks, so there is approximately a 25% chance all three stocks will be listed on the same exchange, and so on.

Given the empirical split of 56:44 mentioned above, we can reject the hypothesis that the majority of investors show a distinct preference for stocks from one exchange. Thus, there appears to be little evidence of other transaction costs that lead investors to invest in one exchange at a time.

6.2 Industry expertise/affinity

Industry expertise/affinity may explain the location of trade bias. Suppose a certain industry tends to list in Shanghai and another industry tends to list in Guangdong. If Shanghai investors work in the first industry, they might believe they have private information about the industry. Or, they might feel more comfortable investing in companies from that industry.

We check the hypothesis that certain industries might be linked to one exchange or another. Appendix 2 shows the distribution of industries for the market as a whole, Shanghai-listed companies, and Guangdong-listed companies. A Wilcoxon sign-ranked test fails to reject the hypothesis that both distributions are the same (results not reported). This finding matches our research. In the first decade of public stock exchanges in the PRC, the China Securities Regulatory Commission (CSRC), the regional CSRCs, and industry ministries worked to match the number, capital, and industry affiliations of companies on the two stock exchanges.

6.3 Non-traded goods

In a world with non-traded goods, investors may have hedging demands that can be met by holding local stocks—see Stockman and Dellas (1989) and Baxter, Jermann, and King (1998). Whether an investor over- or underweights the locally headquartered stock depends on whether the investor's assumed utility function is separable between traded and non-traded goods and the level of risk aversion.

While non-traded goods can possibly explain pure home bias or cultural affinity bias, it is hard to imagine they can explain the location of trade bias that dominates our results. This could only happen if certain types of companies list on one exchange and certain types list on the other. The type of company that lists in Shanghai would have to provide hedging to Shanghai investors and not to Guangdong investors. What's more, investors in the non-exchange regions would have to have hedging demands that were not related to those in either Shanghai and Guangdong. Appendix 2 shows little difference in the industries listed on the two exchanges.

6.4 Community effects

Thus far, we have used the market as the reference portfolio. It is possible that investors do not have mean-variance preferences. For example, investors care about aggregate community wealth as well as their own wealth. DeMarzo, Kaniel, and Kremer (2004) offer such a model. In such cases, we may measure a non-zero over/underweighting of locally headquartered stocks even though the true value might be zero. The error would stem from mis-specifying the reference portfolio.

The DeMarzo, Kaniel, and Kremer (2004) model might also explain the cultural affinity bias if investors want their wealth to exceed that of family and friends back in the region where they were born. Such an explanation makes sense in the modern-day PRC, where people travel to large cities to look for work. Money earned while working in a big city is often sent back to the family or saved until the person returns home. Investors might overweight stocks that are headquartered where the investor was born *relative to the market portfolio.*

Community effects cannot explain overweighting of locally listed stocks. Locally listed stocks are not necessarily headquartered in the region where the investor lives nor are they necessarily headquartered in the region where the investor was born. Since location of trade is the dominant factor, we conclude that community effects do not explain the majority of our results.

6.5 Habit formation

Shore and White (2002) link external habit formation and home bias. In their model "a small group of agents holds primarily domestic securities." Other agents with external habit formation then mimic the domestic bias of the small group. The authors envisage small business owners who are forced to hold local assets for agency reasons. In our case, there is no reason to think that a group of Shanghai investors are forced to hold Beijing stocks that are listed in Shanghai, as opposed to Beijing stocks that are listed in Guangdong.

6.6 Attention

A possible explanation of our findings relates to recent work by Barber and Odean (2003). The authors hypothesize that investors face a difficult search problem when deciding which stocks to buy. In the PRC, investors must choose between almost 1,000 stocks; in the U.S. the number is an order of magnitude larger.

While attention may help explain why Shanghai investors limit themselves to buying Shanghai-listed stocks and Guangdong investors limit themselves to Guangdong-listed stocks, it does not help explain the behavior of investors from the non-exchange regions. As we show above, the majority of these investors hold stocks from both exchanges and do not limit themselves to stocks from one exchange or the other.

6.7 Better diversified investors

One worry in a study of investor behavior is that departure from rationality makes null hypotheses hard to verbalize. For example, does it make sense to use the CAPM portfolio as a reference portfolio if investors hold only three stocks? Clearly these investors care little about diversification. Therefore, we split our sample into those who hold few stocks (less than seven) and those who hold more stocks (seven or more) and redo the results in Table V. Regression estimates can be seen in Appendix 3. It is apparent that better diversified investors exhibit the same portfolio tilting as less diversified investors. Appendix 3 shows only regression coefficients (and not t-stats) for clarity.

6.8 Transformation of our over/underweighting measure

Readers might worry that the left-hand side (LHS) variables in the regressions from Table V and Appendix 3 are bounded (-100% , +100%). These limits might affect either coefficient estimates or significance tests. We redo the results in Table V after transforming the LHS variable. The transformation entails mapping the variable on the (0,1) interval and taking a logit transformation. Results and statistical significance (not reported) are not materially changed. That is, location of trade is still the dominant bias and it is statistically significant.

6.9 Distance from the exchange

Ideally we would like to test gradations of location effects, but we are limited by our sample. Since we have investors from only seven regions, it is not possible to test whether investors who live in a region next to an exchange tend to exhibit a location of trade bias. The data isn't available to us at this time.

6.10 Notes on our cultural affinity measure

It is possible that our cultural affinity measurements are affected by our interpretation of the NIC numbers. When an individual is first given an internal passport, the NIC number is coded with the individual's birth date, place of residence, and gender. If an individual moves, he or she is supposed to re-register with the police in the new location. The re-registration results in a new NIC number that reflects the new place of residence. It is possible that investors in our sample have moved and re-registered. Thus, what we interpret to be an investor who is currently living in the region where he or she was born may actually be an investor who has moved to the region and re-registered.

Our statistical power comes from investors whose birth region is different from their home region. Any mis-specification would lead us to focus on recent or temporary migrants. These are the investors that one would expect to have the strongest cultural affinity bias. Yet the cultural affinity bias is economically and statistically small in our study. Thus, we are confident that the location of trade bias is the dominant effect.

6.11 Correlation of independent variables

We check the correlation of the right-hand side indicator variables used in the regression and ANOVA analyses. Not surprisingly, $Dum_{Home=HQ}$ and $Dum_{Birth=HQ}$ are correlated with a 0.8 coefficient. Individuals in the PRC tend to live and work in the region where they were born. The correlation of $Dum_{Home=Listing}$ with the other two indicator variables ranges between 0.15 and 0.25, depending on whether you look at all investors or only those from Shanghai and Guangdong. This is additional support that we have the statistical power to conclude that location of trade is the dominant factor.

7 Conclusion

This paper studies the portfolio choices of individual investors. We test the implication of a multi-asset rational expectations model when some investors receive better (more) information about a subset of stocks than others receive.

At first glance, holdings appear to fit the model well. Investors overweight stocks where we hypothesize they have better information. For example, investors who live close to a firm tend to hold more of that firm's stock than those who live farther away. We also examine stocks that offer no relative advantage to any partition of investors. Investors allocate the same fraction of their portfolio to the largest 237 stocks as the aggregate market does. Finally, there is weak evidence that the degree of home bias is higher for assets that have a large covariance with the market portfolio.

Further examination leads to a different interpretation of our results. Three series of tests all support the conclusion that investors are more familiar with certain stocks, but do not possess valuable information. Familiarity manifests itself in a predictable way—certain stocks are overweighted in investors' portfolios. We show the fraction of investor's portfolio held in local stocks does not outperform the fraction held in distant stocks. Events (IPOs) that happened many years in the past can predict stock holding levels today—even though the PRC is a high-turnover market.

Our finding that the location of trade bias dominates investor portfolios may be a bit unsettling at first. However, this result are consistent with a number of recent studies of cross-border ADR holdings and makes us reconsider past studies that have focused only on home bias. We reinterpret existing home bias findings as coming (in large part) from a location of trade bias. In other words, investors appear to have a strong preference for locally listed stocks in both international and intra-national markets. In the United States most locally listed stocks are from U.S. companies. Therefore, researchers may have assumed that U.S. investors prefer stocks of U.S. companies when investors actually prefer U.S.-listed stocks. Disentangling these preferences has been difficult up until now, because the U.S. and foreign markets operate in different time zones, in different currencies, with different trading costs, and under different legal systems. Our research design exploits a unique market structure with one legal system, one accounting system, two exchanges, and no cross-listing.

Finally, our results help explain recent studies that link location of trade to differences in asset prices. Studies such as Froot and Dabora (1999) and Chan, Hameed, and Lau (2003) attribute their findings to country-specific investor sentiment. This paper shows that investors are actually segregated by location of trade. This segregations allows for the possibility that different shocks affect different groups of investors differentially (and may affect asset prices differentially.) Clearly, there is a wealth of future projects that can examine how clientele effects in different locations affect asset prices.

References

[1] Ackert, Lucy F., Bryan K. Church, James Tompkins, and Ping Zhang, 2003, "What's in a Name? An Experimental Examination of Investment Behavior," Working paper, Federal Reserve Bank of Atlanta.

[2] Admati, Anat R., 1985, "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, 53, 3, May, 629-657.

[3] Agronin, Eugene, 2003, "A New Measure of Home Bias and its Determinants," Working paper.

[4] Ahearne, Alan G., William L. Griever, and Francis E. Warnock, 2004, "Information Costs and Home Bias: An Analysis of U.S. Holdings of Foreign Equities," *Journal of International Economics,* 62, 2, March, 313-336.

[5] Al-Khail, Mohammed Aba, and Tom Berglund, 2004, "What Explains the Country Allocation of International Portfolio Investments? Universally Available Versus Trade Related Information," Working paper, Swedish School of Economics and Business.

[6] Amadi, Amir, 2004, "Equity Home Bias: A Disappearing Phenomenon?" Working paper, U.C. Davis.

[7] Ammer, John, Sara B. Holland, David C. Smith, and Francis E. Warnock, 2004, "Look at Me Now: The Role of Cross-Listings in Attracting U.S. Shareholders," Working paper, Board of Governors of the Federal Reserve System.

[8] Athanasios, Koulakiotis, Katerina Lyroudi, and Dimitrios Angelidis, 2004, "The Impact of Foreign Cross-Listing on Symmetric Information Spillovers Between Markets: The Case of the Paris Stock Exchange," Working paper, University of Wales.

[9] Baker, H. Kent, John R. Nofsinger, and Daniel G. Weaver, 2002, "International Cross-Listing and Visibility," *Journal of Financial and Quantitative Analysis,* 37, 3, September.

[10] Barber, Brad M. and Terrance Odean, 2000, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance,* LV, 2, April, 773-806.

[11] Barber, Brad M. and Terrance Odean, 2003, "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," Working paper, U.C. Davis, October.

[12] Barber, Brad M., Terrance Odean, and Michal Strahilevitz, 2003, "Once Burned, Twice Shy: Nave Learning, Counterfactuals, and the Repurchase of Stocks Previously Sold," Working paper, U.C. Davis.

[13] Bedi, Jaideep, Anthony Richards, and Paul Tennant, 2003, "The Characteristics and Trading Behavior of Dual-Listed Companies," *Working Paper,* Reserve Bank of Australia.

[14] Bekaert, Geert and Campbell R. Harvey, 2000, "Capital Flows and the Behavior of Emerging Market Equity Returns," In: Edwards, S. (Ed.), Capital Inflows and the Emerging Economies. University of Chicago Press and NBER, Cambridge, MA, 159-194.

[15] Bekaert, Geert, Campbell R. Harvey, and Robin L. Lumsdain, 2002, "The Dynamics of Emerging Market Equity Flows," *Journal of International Money and Finance*, 21, 295-350.

[16] Benartzi, Shlomo and Richard H. Thaler, 2001, "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 1, March, 79-98.

[17] Ben-Haim, Yakov and Karsten Jeske, 2003, "Home Bias in Financial Markets: Robust Satisficing with Info Gaps," Working paper, Federal Reserve Bank of Atlanta.

[18] Bodnaruk, Andriy, undated, "Proximity Always Matters: Evidence from Swedish Data," Working paper, Stockholm School of Economics.

[19] Brennen, Michael J. and H. Henry Cao, 1997, "International Portfolio Investment Flows," *Journal of Finance,* LII, 5, December, 1851-1880.

[20] Brennen, Michael J. and Walter N. Torous, 1999, "Individual Decision Making and Investor Welfare," *Economic Notes by Banca Monte dei Paschi di Siena SpA,* 28, 2, 119-143.

[21] Cai, Fang and Francis E. Warnock, 2004, "International Diversification at Home and Abroad," Working paper, Board of Governors of the Federal Reserve System.

[22] Chan, Kalok, Vincentiu Covrig, and Lilian Ng, 2004, "What Determines the Domestic Bias and Foreign Bias? Evidence from Equity Mutual Fund Allocations Worldwide," Working paper, HKUST.

[23] 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.

[24] Cooper, Ian and Evi Kaplanis, 1994, "Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium," *Review of Financial Studies,* 7, 1, Spring, 45-60.

[25] Coval, Joshua D. and Tobias J. Moskowitz, 1999, "Home Bias at Home: Local Equity Preference in Domestic Portfolio," *Journal of Finance*, LIV, 6, December 2045-2073.

[26] Coval, Joshua D. and Tobias J. Moskowitz, 2001, "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy*, 109, 4, 811-841.

[27] Dahlquist, Magnus and Goran Robertsson, 2001, "Direct Foreign Ownership, Institutional Investors, and Firm Characteristics," *Journal of Financial Economics,* 59, 413-440.

[28] DeMarzo, Peter M., Ron Kaniel, and Ilan Kremer, 2004, "Diversification as a Public Good: Community Effects in Portfolio Choice," *Journal of Fiannce,* Auguest.

[29] Disyatat, Piti and R. Gaston Gelos, 2001, "The Asset Allocation of Emerging Market Mutual Funds," Working paper, International Monetary Fund.

[30] Edison, Hali J. and Francis E. Warnock, 2003, "U.S. Investors' Emerging Market Equity Portfolios: A Security-Level Analsyis," Forthcoming *Review of Economics and Statistics.*

[31] Errunza, Vihang, Ked Hogan, and Mao-Wei Hung, 1999, "Can the Gains from International Diversification Be Achieved Without Trading Abroad?" *Journal of Finance,* LIV, 6, December, 2075-2107.

[32] Feng, Lei and Mark S. Seasholes, 2003, "A Profile of Individual Investors in an Emerging Stock Market," Working paper, U.C. Berkeley.

[33] Feng, Lei and Mark S. Seasholes, 2004, "Correlated Trading and Location," forthcoming *Journal of Finance.*

[34] French, Kenneth R and James M. Poterba, 1991, "Investor Diversification and International Equity Markets," *American Economic Review,* 81, 2, 222-226.

[35] 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.

[36] Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 5, September, 1121-1152.

[37] Goetzmann, William N. and Alok Kumar, 2004, "Diversification Decisions of Individual Investors and Asset Prices," Working paper, Yale University.

[38] Grinblatt, Mark and Matti Keloharju, 2000, "The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set," *Journal of Financial Economics*, 25, 43-67.

[39] Grinblatt, Mark and Matti Keloharju, 2001, "How Distance, Language, and Culture Influence Stockholdings and Trades," *Journal of Finance*, LVI, 3 June, 1053-1073.

[40] Hau, Harald, 2001, "Location Matters: An Examination of Trading Profits," *Journal of Finance*, LVI, 5, October 1959-1983.

[41] Hertz, Ellen, 1998, "The Trading Crowd: An Ethnography of the Shanghai Stock Market," Cambridge University Press.

[42] Hiraki, Takato, Hideaki Inoue, Akitoshi Ito, and Hiroyuki Masuda, 2003, "Investor Biases in Japan," Working paper, International University of Japan.

[43] Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2003, "Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers," Working paper, NBER 9711.

[44] Huberman, Gur, 2001, "Familiarity Breeds Investment," *Review of Financial Studies*, 14, 3, 659-680.

[45] Ivkovic, Zoran and Scott Weisbenner, 2004, "Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments," Forthcoming *Journal of Finance.*

[46] Kang, Jun-Koo and Rene M. Stulz, 1997, "Why Is There a Home Bias? An Analysis of Foreign Portfolio Equity Ownership in Japan," *Journal of Financial Economics,* 46, 3-28.

[47] King, Michael R. and Dan Segal, 2003, "Corporate Governance, International Cross Listing and Home Bias," Working paper, Bank of Canada.

[48] Lease, Ronald C, Wilbur G. Lewellen, and Gary G. Schlarbaum, 1974, “The Individual Investor: Attributes and Attitudes,” *Journal of Finance,* 29, May, 413-433.

[49] Levy, Haim and Marshall Sarnat, 1970, “International Diversification of Investment Portfolios,” *American Economic Review,* 60, 4, 668-75.

[50] Lewis, Karen K., 1995, “Puzzles in International Financial Markets,” in Handbook of International Economics, Grossman and Rogoff ed. Amsterdam: North Holland, 1913-1917.

[51] Lewis, Karen K., 1996, “What Can Explain the Apparent Lack of International Consumption Risk Sharing?” *Journal of Political Economy*, 104, 21, 267-297.

[52] Lewis, Karen K., 1999, “Trying to Explain Home Bias in Equities and Consumption,” *Journal of Economic Literature*, XXXVII, June, 571-608.

[53] Linnainmaa, Juhani, 2003, “The Anatomy of Day Traders,” Working paper, UCLA.

[54] Lintner, John, 1965, “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economic and Statistics,* 13-37.

[55] Lutje, Torben and Lukas Menkhoff, 2004, “What Drives Home Bias? Evidence from Fund Managers’ Views,” Working paper, University of Hannover.

[56] Massa, Massimo and Andrei Simonov, undated, “Behavioral Biases and Portfolio Choice,” Working paper, INSEAD.

[57] Merton, Robert C., 1987, “A Simple Model of Capital Market Equilibrium with Incomplete Information,” *Journal of Finance,* XLII, 3, July, 483-510.

[58] Morck, Randall, Bernard Yeung, and Wayne Yu, 2000, “The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements?,” *Journal of Financial Economics,* 58, 215-260.

[59] Morse, Adair and Sophie Shive, 2003, “Patriotism in Your Portfolio,” Working paper, University of Michigan.

[60] Sharpe, William F., 1964, “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk,” XIX, 3, September, 425-442.

[61] Shleifer, Andrei and Robert W. Vishy, 1997, “The Limits of Arbitrage,” *Journal of Finance,* LII, 1, March.

[62] Shore, Stephen H. and Joshua S. White, 2002, "External Habit Formation and the Home Bias Puzzle," Working paper, Harvard University.

[63] Solnik, Bruno H., 1974, "The International Pricing of Risk: An Empirical Investigation of the World Capital Market Structure," *Journal of Finance,* 29, 2, 365-378.

[64] Statman, Meir, 1987, "How Many Stocks Make a Diversified Portfolio?" *Journal of Financial and Quantitative Analysis,* 22, 3, September, 353-363.

[65] Statman, Meir, 2002, "How Much Diversification is Enough?," Working paper, Santa Clara University.

[66] Stockman, Alan C. and Harris Dellas, 1989, "International Portfolio Nondiversification and Exchange Rate Variability," *Journal of International Economics,* 26, 3 / 4, 271-289.

[67] Stulz, Rene M., 1981, "On the Effects of Barriers to International Investment," *Journal of Finance,* XXXVI, 4, September, 923-934.

[68] Tesar,Linda L. and Ingrid M. Werner, 1995, "Home Bias and High Turnover," *Journal of International Money and Finance,* 14, 4, 467-492.

[69] Uppal, Ramman and Tan Wang, 2003, "Model Misspecification and Underdiversification," forthcoming *Journal of Finance.*

[70] Weitzman, Martin L., 1979, "Optimal Search for the Best Alternative," *Econometrica,* 47, 3, 641-654.

[71] Zhu, Ning, 2003, "The Local Bias of Individual Investors," Working paper, Yale University.

Figure 1
People's Republic of China: Provinces, Autonomous Regions, and Municipalities

11. Beijing
12. Tianjin
13. Hebei
14. Shanxi
15. Neimenggu (Inner Mongolia)
21. Liaoning
22. Jilin
23. Heilongjiang
31. Shanghai
32. Jiangsu
33. Zhejiang
34. Anhui
35. Fujian
36. Jiangxi
37. Shandong
41. Henan
42. Hubei
43. Hunan
44. Guangdong (Canton)
45. Guangxi
46. Hainan
50. Chongqing
51. Sichuan
52. Guizhou
53. Yunnan
54. Xiazang (Tibet)
61. Shaanxi
62. Gansu
63. Qinghai
64. Ningxia
65. Xinjiang

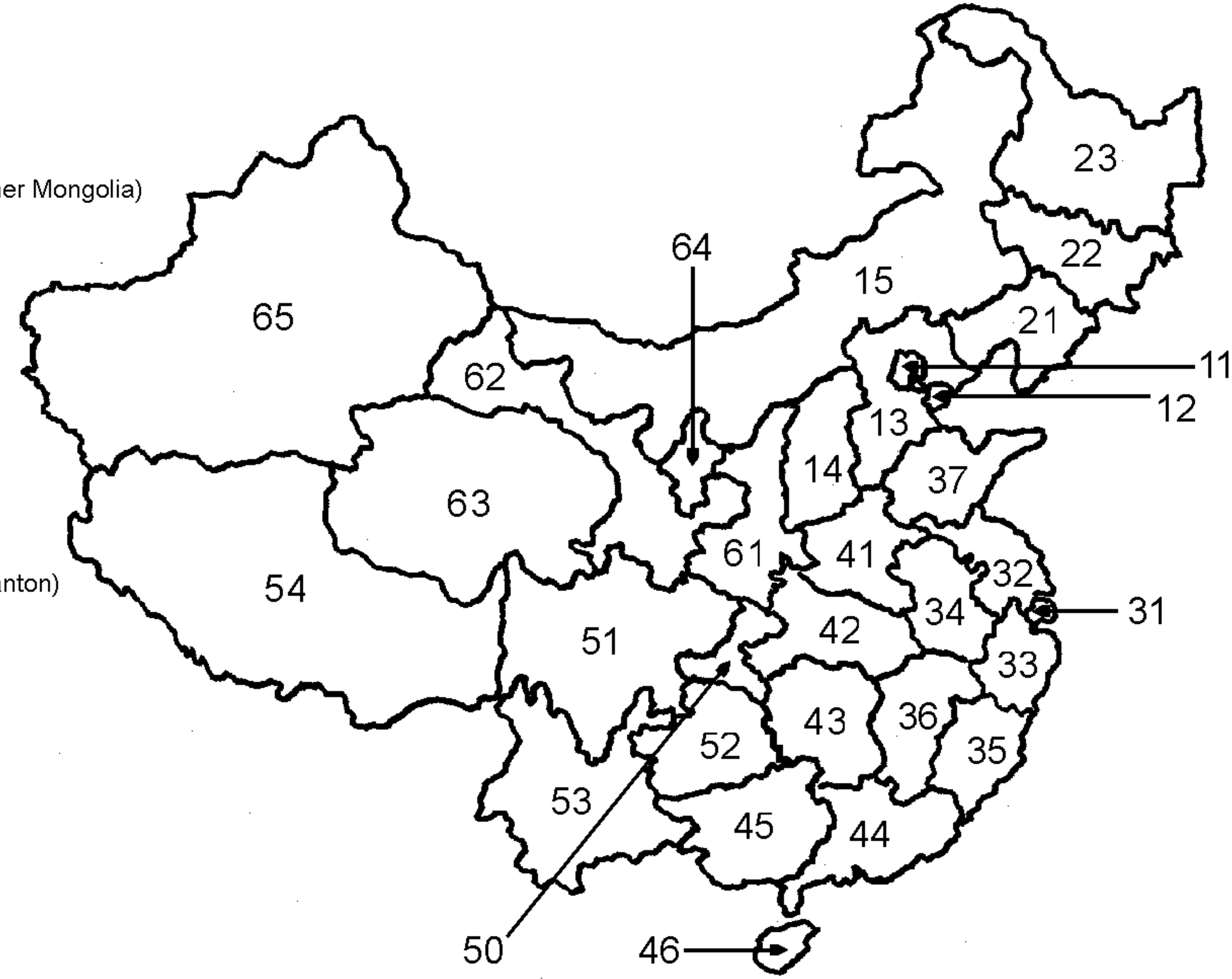

Table I
Testable Hypotheses and Information Advantages

This table presents testable hypotheses of stock holdings and information advantages. Panel A lists stocks where a partition of investors is hypothesized to have an information advantage for a partition of stocks. Panel B presents stocks where there is no apparent partition of investors with an informational advantage. Stocks that offer an information advantage to a partition of investors should be over-weighted in those investors' portfolios and under-weighted in all other investors' portfolios. The fraction of investors' portfolios allocated to stocks without clear advantages should be equal across any partition of investors.

Panel A: Partitions of stocks and investors where an information advantage may exist

Testable hypothesis	Stock partition	Investor partition
A.1	firm is from a given geographic region	lives near firm
A.2	firm is from a given cultural region	is from same culture
A.3	stock listed on a given exchange	lives near exchange
A.4	CEO speaks a given language	speaks same language
A.5	firm publishes reports in a given language	reads same language
A.6	firm is in a given sector	works in the same sector
A.7	firm produces specialized/localized products	lives near firm
hard to monitor firms		
A.8	firm has small market capitalization	lives near firm
A.9	firm is highly levered	lives near firm
A.10	firm has a high fraction of intangible assets	lives near firm
A.11	firm has especially poor financial records	lives near firm

Panel B: Partitions of stocks and investors where an information advantage may not exist

Testable hypothesis	Stock partition	Investor partition
B.1	firm has large market capitalization	any partition
B.2	firm produces widely-distributed products	any partition
B.3	firm is part of well known index	any partition

Table II
Portfolio Overview Statistics

The table presents some general overview statistics of the stock holding data. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000. Other time periods from 1999 to 2000 are used as control groups. We have 51,218 distinct individual accounts that hold over RMB 7 bn. At an exchange rate of RMB 8 : USD 1 the holdings are approximately USD 1 bn.

(1) Location of branch (region in PRC)	(2) Number of branch offices (#)	(3) Number of investors (#)	(4) Median number of stocks per account (#)	(5) Average portfolio value (RMB)	(6) Median portfolio value (RMB)	(7) Total portfolio value held by all investors (RMB)
Beijing	1	7,604	3	134,209	34,745	1,020,521,664
Guangdong	4	6,488	2	273,194	53,105	1,772,480,640
Heilongjiang	1	7,408	2	48,187	22,889	356,971,872
Hubei	1	4,399	2	90,893	32,455	399,836,800
Shandong	1	5,299	2	78,145	26,650	414,091,520
Shanghai	5	14,559	3	162,069	44,060	2,359,566,848
Sichuan	2	5,461	2	124,878	28,752	681,957,248
Total	15	51,218	--	--	--	7,005,426,592
Average	--	--	3	136,777	34,442	--

Table III
Preliminary Analysis: Information Advantage May Exist

The table overviews portfolio holdings for three partitions of investors and listed stocks. Our data come from fifteen branch offices of one brokerage firm. We have 51,218 distinct individual accounts that hold over RMB 7 bn. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000.

Panel A: Pure Home Bias (*hypothesis A.1*)

Region in PRC	Average fraction of portfolio held in the region where investor currently lives	Market capitalization of firms in region as a fraction of total market cap	Average overweight due to pure home bias
Beijing	0.1042	0.0671	0.0371
Guangdong	0.3131	0.2339	0.0792
Heilongjiang	0.0691	0.0200	0.0491
Hubei	0.1067	0.0339	0.0728
Shandong	0.0351	0.0090	0.0261
Shanghai	0.3724	0.2189	0.1535
Sichuan	0.1077	0.0399	0.0678
Average	0.1953	0.1128	0.0825

Panel B: Cultural Affinity Bias (*hypothesis A.2*)

	Average fraction of portfolio from region where investor was born	Average fraction of market from region where investor was born	Average overweight due to cultural affinity
Average	0.1692	0.0933	0.0758

Panel C: Location of Trade Bias (*hypothesis A.3*)

Region in PRC	Fraction of investor's portfolio held in locally listed stocks	Fraction of PRC market cap listed on local exchange	Average overweighting due to location of trade
Guangdong	0.8228	0.5157	0.3071
Shanghai	0.8431	0.4843	0.3588
Average	0.8368	0.4940	0.3428

Table IV
Preliminary Analysis: Information Advantage May Not Exist

The table overviews portfolio holdings. Panels A, B, and C consider the largest 237 stocks—those in the top 25% when ranked by market capitalization. Investors are partitioning three different way: i) as a whole group; ii) by province where they currently live; and iii) and by branch office where they opened their account. Panel D partitions stock by the exchange where they trade and considers the holdings of investors who currently live in the same region. Our data come from fifteen branch offices of one brokerage firm. We have 51,218 distinct individual accounts that hold over RMB 7 bn. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000.

Panel A: Weight of largest 237stocks
all investors in our sample (*hypothesis B.1*)

Fraction of portfolio allocated to the 237 largest stocks	Weight of the 237 largest stocks in overall market	overweighting
0.5813	0.5794	0.0019

Panel B: Weight of largest 237 stocks
partitioned by investor province (*hypothesis B.1*)

Investor province	Fraction of portfolio allocated to the 237 largest stocks	Weight of the 237 largest stocks in overall market	overweighting
Beijing	0.5925	0.5794	0.0131
Guangdong	0.6259	0.5794	0.0465
Heilongjiang	0.5537	0.5794	-0.0257
Hubei	0.6167	0.5794	0.0373
Shandong	0.5632	0.5794	-0.0162
Shanghai	0.5369	0.5794	-0.0425
Sichuan	0.6070	0.5794	0.0276
Average	0.5813	0.5794	0.0019

Panel C: Weight of largest 237stocks
partitioned by branch office (*hypothesis B.1*)

average overweighting	median overweighting	inter-quartile range of overweighting
0.0019	0.0010	[-0.0383 , +0.0691]

Table V
Regression Results of Portfolio Tilting (Biases)

The table shows results for three distinct biases in a regression framework. For each investor, we calculate the percentage (weight) of his or her portfolio that is invested in each region/exchange combination (or "bin"). We calculate the percentage of the market in each bin. We regress the difference of these two weights on three indicator (dummy) variables. Regressions 1, 2, and 3 repeat the results from Table II. Regression 4 tests for all three biases simultaneously. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000. Regressions use all 3,175,516 investor/region/exchange combinations. T-statistics (in parentheses) are based on robust standard errors and correct for the actual number of non-zero investor/region/exchange combinations in our data set (149,691). Statistical significance and fit are discussed more thoroughly in the text.

$$\omega_{r,e}^{m} - \omega_{r,e}^{*} = \gamma_1 Dum_{r,e}^{Home=HQ} + \gamma_2 Dum_{r,e}^{Born=HQ} + \gamma_3 Dum_{r,e}^{Home=Listing} + \varepsilon_{r,e}^{m}$$

		Reg. 1	Reg. 2	Reg. 3	Reg. 4
Investor currently lives in same region as company's headquarters (pure home bias)	$Dum_{Home=HQ}$	0.0825 *(13.56)*			0.0649 *(5.86)*
Investor was born in same region as company's headquarters (cultural affinity)	$Dum_{Born=HQ}$		0.0758 *(13.34)*		0.0156 *(1.52)*
Investor currently lives in region where company's stock is listed (location of trade)	$Dum_{Home=Listing}$			0.3428 *(16.89)*	0.3026 *(15.88)*

Table VI
ANOVA Results of Portfolio Tilting (Biases)

The table shows results for three distinct biases in an ANOVA framework. For each investor, we calculate the percentage (weight) of his or her portfolio that is invested in each region/exchange combination (or "bin"). We calculate the percentage of the market in each bin. We regress the difference of these two weights on a constant and three indicator (dummy) variables. Of the three indicator variables, the location of trade bias explains the highest percentage of variance. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000.

$$\omega_{r,e}^{m} - \omega_{r,e}^{*} = \gamma_1 Dum_{r,e}^{Home=HQ} + \gamma_2 Dum_{r,e}^{Born=HQ} + \gamma_3 Dum_{r,e}^{Home=Listing} + \varepsilon_{r,e}^{m}$$

	Partial SS	df	MS	F	Prob > F
Model	282.93	3	94.31	10,649	0.0000
$Dum_{Home=HQ}$	30.12	1	30.12	3,401	0.0000
$Dum_{Born=HQ}$	2.38	1	2.38	268	0.0000
$Dum_{Home=Listing}$	100.58	1	100.58	11,353	0.0000

Table VII
Multivariate Factor Model of Portfolio Tilting (Biases)

The table shows results for three distinct biases with a multivariate factor model. For each investor, we calculate weekly returns to his or her portfolio, over the next half year, based on holdings at one point in time. Individual returns are grouped into portfolios based on location. We describe the procedure for forming investor portfolios in the text. We also calculate the returns of the market portfolio, exchange portfolios, and regional portfolios. We regress the returns from each portfolio (of investor returns) on these factors. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC). T-statistics (in parentheses) are based on robust standard errors. Sample size and fit are discussed in the text.

$$r_t^m = \alpha + \beta \cdot r_t^{mkt} + \Gamma_1 \cdot r_t^{\mathrm{home}} + \Gamma_2 \cdot r_t^{born} + \Gamma_3 \cdot r_t^{exch} + \varepsilon_t^m$$

	Reg. 1	Reg. 2	Reg. 3	Reg. 4a	Reg. 4b
Sample	all	all	all	All investors except for Guangdong and Shanghai	Guangdong and Shanghai investors only
r_{mkt}	0.9895 *(41.98)*	0.9086 *(39.39)*	0.9544 *(85.15)*	0.9920 *(41.66)*	0.5180 *(3.03)*
$r_{home\ region}$		0.0798 *(3.95)*			
$r_{birth\ region}$			0.0465 *(4.75)*		
$r_{local\ exch.}$					0.4768 *(2.81)*
Constant	-0.0009 *(-1.06)*	-0.0009 *(-2.73)*	-0.0010 *(-5.45)*	-0.0010 *(-1.15)*	-0.0086 *(-1.38)*

Table VIII
Economic Costs of Portfolio Tilting (Biases)

The table provides a rough estimate of the economic costs of underdiversification. To simplify our calculations, we use weekly Sharpe ratios. We calculate the Sharpe ratio an investor would have achieved if s/he had held the market portfolio or one of the two exchange portfolios. We also calculate the average Sharpe ratio to holding one of the thirty-one provincial portfolios. Finally, we calculate the average Sharpe ratios an investor would have achieved from holding a portfolio with fifty stocks or a portfolio with only five stocks. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time (01-Jun-2000.)

(1)	(2) Sharpe Ratio	(3) "Cost" (difference from market)	(4) Average over/underweighting (from Table V)
Holds entire market	0.1356	0.0000	0.0000
Holds stocks from one exchange (only)	0.1253	0.0103	0.3026
Holds stocks from one region (only)	0.1054	0.0302	0.0649
Holds 5 stocks (only)	0.0801	0.0555	
Average investor in our sample	0.1198	0.0158	

Table IX
Performance of Local vs. Non-Local Stocks

We test whether the local portion of an investor's portfolio outperforms the non-local portion. For each investor in our sample (*N=51,218*) we have the portfolio holdings as of 01-Jun-2000. We then calculate the gross returns over the next six months. We regress the gross returns on the percent of the portfolio held in local stocks and the percent held in non-local stocks using three different measures of locality. For each regression we provide six different measures of the standard error of the difference. The last four measures allow for clustering of data at different levels. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC).

$$R_{t+k}^{m} = \beta_{L}\left(\%L_{t}^{m}\right) + \beta_{NL}\left(\%NL_{t}^{m}\right) + \varepsilon_{t}^{m}$$

Panel A: Pure home bias

	β_L coef. on share of local stocks	β_{NL} coef. on share of non-local stocks	β_L-β_{NL} locality difference
Coefficient estimate	0.1164	0.1288	-0.0124
OLS standard errors			*(0.0024)*
Robust standard errors			*(0.0026)*
clustering at major holding level (941 groups)			*(0.0096)*
clustering at birth-home level (190 groups)			*(0.0084)*
clustering at birth level (31 groups)			*(0.0081)*
clustering at home level (7 groups)			*(0.0115)*

Panel B: Cultural affinity bias

	β_L coef. on share of local stocks	β_{NL} coef. on share of non-local stocks	β_L-β_{NL} locality difference
Coefficient estimate	0.1205	0.1276	-0.0070
OLS standard errors			*(0.0025)*
Robust standard errors			*(0.0027)*
clustering at major holding level (941 groups)			*(0.0091)*
clustering at birth-home level (190 groups)			*(0.0096)*
clustering at birth level (31 groups)			*(0.0103)*
clustering at home level (7 groups)			*(0.0096)*

Panel C: Location of trade bias

	β_L coef. on share of local stocks	β_{NL} coef. on share of non-local stocks	β_L-β_{NL} locality difference
Coefficient estimate	0.1310	0.1239	0.0071
OLS standard errors			*(0.0016)*
Robust standard errors			*(0.0017)*
clustering at major holding level (941 groups)			*(0.0040)*
clustering at birth-home level (190 groups)			*(0.0065)*
clustering at birth level (31 groups)			*(0.0064)*
clustering at home level (7 groups)			*(0.0071)*

Table X
Information Structure and Long-Lived Effects

We test whether past differences in information are correlated with current holdings. We obtain a proprietary database covering initial public offerings (IPOs). For each of the 778 companies in the database, we calculate the total shares held by investors in our sample on 01-Jun-2000. Our data come from fifteen branch offices of one brokerage firm. The branch offices are located in seven regions in the People's Republic of China (PRC).

Panel A: Pure Home Bias

We test whether investors who live in the same region as a company's headquarters are more likely to hold the company's shares if the investment who managed the initial public offering (IPO) is headquartered in the same region. For each of the 778 companies in the database, we calculate the fraction of all shares in our sample held by investors who live in the same region as the company's headquarters. We call this fraction, (Ω). For the 388 companies with at least one local investor in our sample, we regress the measure on a constant and an indicator variable. The indicator variable equals one if the investment bank is headquartered in the same region as the company. T-statistics are based on robust standard errors.

$$\Omega_i^{pure\,home\,bias} = \alpha + \gamma \cdot Dum_i^{I-bank} + \varepsilon_i$$

α	γ
0.2652	0.1893
(14.38)	*(7.61)*

Panel B: Cultural Affinity Bias

We test whether investors who were born in the same region as a company's headquarters are more likely to hold the company's shares if the investment who managed the initial public offering (IPO) is headquartered in the same region. For each of the 778 companies in the database, we calculate the fraction of all shares in our sample held by investors who live in the same region as the company's headquarters. We call this fraction, (Ω). For the 595 companies with at least one local investor in our sample, we regress the measure on a constant and an indicator variable. The indicator variable equals one if the investment bank is headquartered in the same region as the company. T-statistics are based on robust standard errors.

$$\Omega_i^{cultural\ affinity} = \alpha + \gamma \cdot Dum_i^{I-bank} + \varepsilon_i$$

α	γ
0.1117	0.1652
(13.00)	*(10.32)*

Panel C: Location of Trade Bias

We test whether investors who were live in the same region where a company's stock is listed are more likely to hold the company's shares if the investment who managed the initial public offering (IPO) is headquartered in the same region. For each of the 778 companies in the database, we calculate the fraction of all shares in our sample held by investors who live in the same region as the company's headquarters. We call this fraction, (Ω). For the 667 companies with at least one local investor in our sample, we regress the measure on a constant and an indicator variable. The indicator variable equals one if the investment bank is headquartered in the region where the company is listed. T-statistics are based on robust standard errors.

$$\Omega_i^{location\ of\ trade} = \alpha + \gamma \cdot Dum_i^{I-bank} + \varepsilon_i$$

α	γ
0.1689	0.2735
(20.35)	*(18.72)*

Appendix 1
Regional Statistics from the PRC

We present a list of the thirty-one regions in the PRC. Column 2 has the same regional code used in Figure 1. Column 3 has the size (in km^2) of the province or municipality as provided by the central government of the People's Republic of China (PRC). Regional GDP and GDP per capita in Columns 4 and 5 are from the central government. Regional population in Column 6 is from the brokerage firm that supplied our data. Column 7 shows monthly household income and is from a private survey.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Region Name	Regional code	Area (km^2)	GDP (RMB bn)	GDP per capita (RMB)	Population (mm)	Monthly household income (RMB)
Anhui	34	139,400	290.9	4,707	66,117,241	636.7
Beijing	11	16,800	217.5	19,846	11,061,983	1,184.2
Chongqing	50	N/A	148.0	4,826	30,723,399	747.3
Fujian	35	136,000	333.0	10,797	32,835,978	1,067.9
Gansu	62	454,000	93.2	3,668	25,074,457	622.5
Guangdong	44	170,000	846.4	11,728	72,988,849	1,336.6
Guangxi	45	236,660	195.3	4,148	46,575,918	860.6
Guizhou	52	176,100	91.2	2,475	36,300,720	623.1
Hainan	46	34,000	47.1	6,383	7,431,864	788.8
Hebei	13	187,700	456.9	6,932	70,102,861	633.1
Heilongjiang	23	453,900	289.7	7,660	36,608,425	490.1
Henan	41	167,000	457.6	4,894	125,809,220	599.6
Hubei	42	187,400	385.8	6,514	59,425,019	754.0
Hunan	43	211,800	332.7	5,105	65,205,272	841.9
Jiangsu	32	102,600	769.8	10,665	70,090,824	950.4
Jiangxi	36	166,600	196.3	4,661	48,538,550	677.2
Jilin	22	187,400	167.0	6,341	26,161,000	548.8
Liaoning	21	144,900	417.2	10,086	42,988,207	617.7
Neimenggu	15	1,183,000	126.8	5,350	23,295,364	583.5
Ningxia	64	51,800	24.2	4,473	5,432,891	636.8
Qinghai	63	791,200	23.8	4,662	4,732,420	610.8
Shaanxi	61	205,600	148.8	4,101	39,097,650	720.5
Shandong	37	153,800	766.2	8,673	89,216,648	794.2
Shanghai	31	6,340	403.5	30,805	13,131,204	1,422.1
Shanxi	14	156,000	150.7	4,727	31,450,808	588.9
Sichuan	51	570,000	371.2	4,452	83,585,559	721.8
Tianjin	12	11,300	145.0	15,976	9,161,665	896.9
Xinjiang	65	1,660,000	217.5	6,470	17,633,656	820.2
Xizang	54	1,201,000	10.6	4,262	2,477,195	N/A
Yunnan	53	394,000	185.6	4,452	40,183,888	747.9
Zhejiang	33	101,800	536.5	12,037	45,123,435	1,456.6
Average				6,919		788.2

Appendix 2
Industry Overview

The table shows the breakdown of companies at the letter level (roughest level) for the entire sample and divided by listing exchange. In the People's Republic of China (PRC), industries classified by one letter and up to four numbers. There are two stock exchanges in the PRC. One is in Shanghai; one is in Guangdong (in the city of Shenzhen). Stocks may not be cross-listed. Industry data provided by brokerage firm in PRC.

Entire Sample

Industry Code	Count	Fraction
0	1	0.11
A	19	2.01
B	10	1.06
C	523	55.34
D	33	3.49
E	15	1.59
F	30	3.17
G	55	5.82
H	82	8.68
I	6	0.63
J	30	3.17
K	34	3.60
L	10	1.06
M	77	8.15
blank	20	2.12
TOTAL	945	

Shanghai-Listed

Industry Code	Count	Fraction
0	0	0.00
A	11	2.27
B	3	0.62
C	252	51.96
D	18	3.71
E	8	1.65
F	18	3.71
G	29	5.98
H	52	10.72
I	3	0.62
J	13	2.68
K	18	3.71
L	7	1.44
M	44	9.07
blank	9	1.86
	485	

Guangdong-Listed

Industry Code	Count	Fraction
0	1	0.22
A	8	1.74
B	7	1.52
C	271	58.91
D	15	3.26
E	7	1.52
F	12	2.61
G	26	5.65
H	30	6.52
I	3	0.65
J	17	3.70
K	16	3.48
L	3	0.65
M	33	7.17
blank	11	2.39
	460	

Appendix 3
Regression Results

We repeat the results shown in Table III, except that we divide our sample into two groups. Group 1 is made up of investors who hold less than seven stocks. Group 2 is made up of investors who hold seven or more stocks. Our data come from fifteen branch offices of one brokerage firm, located in seven regions in the People's Republic of China (PRC). We concentrate on holdings from a single point in time: 01-Jun-2000. We present only coefficient estimates for comparison with Table III.

$$\omega_{i,r,e} - \omega^{*}_{r,e} = \gamma_1 Dum_{Home=HQ} + \gamma_2 Dum_{Born=HQ} + \gamma_3 Dum_{Home=Listing} + \varepsilon_{i,r,e}$$

		# of stocks held	Reg. 1	Reg. 2	Reg. 3	Reg. 4
Investor currently lives in same region as company's headquarters (pure home bias)	$Dum_{Home=HQ}$	< 7	0.0807			0.0631
		≥ 7	0.0960			0.0797
Investor was born in same region as company's headquarters (cultural affinity)	$Dum_{Born=HQ}$	< 7		0.0742		0.0157
		≥ 7		0.0880		0.0136
Investor currently lives in region where company's stock is listed (location of trade)	$Dum_{Home=Listing}$	< 7			0.3522	0.3127
		≥ 7			0.2934	0.2468