THE JOURNAL OF FINANCE • VOL. LX, NO. 6 • DECEMBER 2005

# Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers

HARRISON HONG, JEFFREY D. KUBIK, and JEREMY C. STEIN*

**ABSTRACT**

A mutual fund manager is more likely to buy (or sell) a particular stock in any quarter if other managers in the same city are buying (or selling) that same stock. This pattern shows up even when the fund manager and the stock in question are located far apart, so it is distinct from anything having to do with local preference. The evidence can be interpreted in terms of an epidemic model in which investors spread information about stocks to one another by word of mouth.

IN THIS PAPER, WE EXPLORE THE HYPOTHESIS that investors spread information and ideas about stocks to one another directly, through word-of-mouth communication. This hypothesis comes up frequently in informal accounts of the behavior of the stock market.[1] For example, in his bestseller *Irrational Exuberance*, Shiller (2000) devotes an entire chapter to the subject of "Herd Behavior and Epidemics," and writes

> A fundamental observation about human society is that people who communicate regularly with one another think similarly. There is at any place and in any time a *Zeitgeist*, a spirit of the times. . . . Word-of-mouth transmission of ideas appears to be an important contributor to day-to-day or hour-to-hour stock market fluctuations. (pp. 148, 155)

However, in spite of its familiarity, this hypothesis about word-of-mouth information transmission has received little direct support in stock market data.[2]

*Hong is from Princeton University, Kubik is from Syracuse University, and Stein is from both Harvard University and the National Bureau of Economic Research. Thanks to the National Science Foundation for research support, and to Rebecca Brunson and Ravi Pillai for research assistance. We are also grateful for comments and suggestions from Julian Abdey, Malcolm Baker, Gene D'Avolio, Chip Fortson, Rick Green, Rafael LaPorta, Karl Lins, Burton Malkiel, Anna Scherbina, Andrei Shleifer, Jeff Wurgler, and the referee, as well as from seminar participants at Harvard Business School, Boston College, the University of Texas, New York University, Columbia, Northwestern, Maryland, the University of Southern California, Penn State, Syracuse, and the Western Finance Association meetings.

[1] See, for example, Ellison and Fudenberg (1995) for a formal model of word-of-mouth learning.

[2] However, recent work has done much to advance the more general proposition that local peer group effects can have important consequences for a number of other economic outcomes, including educational attainment and participation in crime. See, for example, Case and Katz (1991), Glaeser, Sacerdote, and Scheinkman (1996), and Bertrand, Luttmer, and Mullainathan (2000); Glaeser and Scheinkman (2002) provide a survey with more references. Relatedly, Dumais, Ellison, and Glaeser

One noteworthy exception is Shiller and Pound's (1989) survey of 131 individual investors. When asked about what had drawn their attention to the stock they had most recently purchased, many of these investors named a personal contact, such as a friend or relative. Although this evidence is certainly suggestive, it seems fair to say that the empirical case for the importance of word-of-mouth communication in the stock market has not yet been fully developed.[3]

In an effort to bring large-sample evidence to bear on the question, we study the holdings and trades of mutual fund managers. Our empirical strategy is a simple one, premised on the assumption that fund managers who work in the same city are more likely to come into direct contact with one another—for example, at local investor conferences and other similar events—and hence to exchange ideas by word of mouth. Consider, for example, a fund manager working for the Fidelity fund family, which is located in Boston, and her decision of whether or not to buy shares of Intel in a given quarter. Our basic prediction is that this decision will be more heavily influenced by the decisions of fund managers working for the Putnam family, which is also located in Boston, than by the decisions of fund managers located in other cities.

This prediction is strongly confirmed in the data. The effects show up when we look at either the level of a fund manager's shareholdings in any given quarter, or the changes in her holdings from one quarter to the next—that is, her trades. Our specifications for trades suggest that controlling for the behavior of the rest of the mutual fund universe, a given manager's purchases of a stock (as a fraction of her total portfolio) increase by roughly 0.13 percentage points when other managers from different fund families in the same city increase their purchases of the same stock by 1 percentage point. So, if managers at Putnam and other Boston-based fund families raise their average weighting on Intel from 1.0% of their portfolios to 2.0%, while other managers across the country keep their weightings fixed, we would expect a typical Fidelity manager to raise her weighting from 1.0% to about 1.13%.

While our analysis is perhaps most naturally motivated by the idea of direct word-of-mouth communication among investors, there are several other possible interpretations of our results, some of which are hard to rule out completely. For example, it might be that mutual fund managers in a given city read the local newspaper or watch a local TV station, and derive some of their investment ideas from this common source. While we cannot reject the possibility that this mechanism is partially responsible for our findings, we would argue that it is ultimately very close in spirit to the word-of-mouth channel, and interesting for many of the same reasons: In both cases, investors receive their information

(2002) and Rosenthal and Strange (2001) provide evidence that geographic proximity facilitates intellectual spillovers across a variety of industries.

[3] There are a few papers that look at word-of-mouth effects in other financial market settings. Kelly and Grada (2000) find striking evidence that during banking panics, bad news about banks spreads via word of mouth in neighborhoods. See also Duflo and Saez (2002) and Madrian and Shea (2000), who demonstrate that individuals' choices of retirement plans are influenced by the decisions of their coworkers.

from somebody who is physically proximate—either a fellow investor, or a local business columnist who acts as a "town crier."

Another alternative story, which we call the "local-investor-relations (LIR) hypothesis," is that fund managers get fed inside information directly by the managers of the companies in which they invest, and that this privileged information distribution happens at the city level.[4] For example, Intel may send a top executive on a road trip to Boston to brief Boston-area fund managers, either at a local conference or in a series of one-on-one office visits. We think it is more worthwhile (and more feasible) to try to discriminate against this hypothesis, and we take steps in this direction.

One piece of evidence against the LIR hypothesis is that we obtain similar results both before and after the implementation of Regulation Fair Disclosure (commonly referred to as "Reg FD"), a set of disclosure rules approved by the Securities and Exchange Commission in the third quarter of 2000 that requires a company to reveal any material information to all investors and analysts simultaneously. The intent of Reg FD is to prohibit exactly the kind of privileged information distribution that is envisioned in the LIR hypothesis. Thus, to the extent that the regulation has any teeth, the fact that our results are not much changed by its implementation suggests that they are not merely the product of an LIR effect. In a similar vein, our results carry over to very small firms with more retail-skewed shareholder bases and scanty analyst coverage, firms that presumably do less in the way of institutionally targeted investor relations. Although we do not claim to be able to completely dismiss the LIR hypothesis, it is again important to bear in mind that for some purposes, our findings are of interest whether or not they are partially driven by an LIR effect. In particular, they help put more structure on the process of information diffusion across investors, suggesting that geography is a key determinant of who learns what when.

An important caveat is that in spite of the colorful quote from Shiller (2000) in the opening paragraph, our approach does not allow us to determine whether fund managers are passing along "irrationally exuberant" sentiment to their nearby colleagues, or real information about fundamentals. Since we only look at data on managers' holdings and trades, and do not show how these link up with stock returns, there is nothing in our evidence that directly contradicts the hypothesis that the market is efficient, and that there are simply frictions associated with gathering information, as in Grossman and Stiglitz (1980).

Our approach to investor behavior in this paper builds on several earlier works. There is a close connection to Coval and Moskowitz (1999, 2001), who also study mutual fund managers, and with whom we share the basic theme

[4] Yet another story—which we take up in more detail below—is that there is reputational herding of the sort described by Scharfstein and Stein (1990), but at the city level. Note that the empirical literature on herding (e.g., Chevalier and Ellison (1999), Hong, Kubik, and Solomon (2000)) has not considered this possibility; instead it has focused on reputational forces that operate at the national level.

that physical proximity facilitates information transmission.[5] The key distinction, however, is that our interest is primarily in the information that investors who are close together can share with one another, whereas Coval and Moskowitz focus on the information that investors can gather about nearby companies. Our own previous work (Hong, Kubik, and Stein (2004)) explores some of the consequences of social interaction for investment behavior, documenting the fact that households that are more socially active are also more likely to invest in the stock market. Here, we analyze professional investors as opposed to households, a setting with many more dollars at stake. Another advantage of the mutual fund data is that we can see funds' decisions with respect to each of thousands of securities, which gives us a great deal of statistical power; in our previous paper, we are limited to examining a household's single binary decision of whether or not to buy stocks at all.

The remainder of the paper is organized as follows. In Section I, we describe our data, as well as the econometric methodology we use. In Section II, we present our main results. Section III reviews some alternative interpretations of these results, and discusses our efforts to discriminate among them. Section IV concludes.

## I. Data and Methodology

Our data on mutual fund holdings come from CDA Spectrum. We augment these data in two ways. First, we associate each mutual fund with a fund family, and use Nelson's Directory of Investment Managers to establish the city in which the fund family is headquartered as of 1998. Second, for each stock, we use the Disclosure database to find the location of the company's headquarters.

In what follows, we analyze both funds' holdings and their trading behavior, with primary emphasis on the trade-based specifications. We begin with eight quarterly snapshots on holdings over the interval from March 1997 to December 1998. The number of mutual funds that make it into our final sample is considerably smaller than the number of funds listed in the Spectrum database over this period, for several reasons. First, we only retain those funds whose headquarters are in the United States, and for whom we are able to find a family location in the Nelson's Directory. Second, we only consider those funds that are classified as stock funds.[6] Third, we eliminate index funds, since their behavior is purely mechanically determined and cannot be influenced by word of mouth.[7] Fourth, not all of the mutual funds in the Spectrum database file holdings reports quarterly; approximately 24% of them only file semiannually

[5] Chen et al. (2004) and Christoffersen and Sarkissian (2002) also look at how mutual funds' locations affect their behavior. Petersen and Rajan (2002) and Berger et al. (2005) are other papers that emphasize a similar theme, but in the context of commercial banks' lending to small companies.

[6] Stock funds are those in the Spectrum database with investment objective codes of 2, 3, 4, or 7, which correspond to aggressive-growth, growth, growth-and-income, and balanced funds, respectively.

[7] We define an index fund as any fund whose name contains the word "index," or some variant thereof. Our results are similar if we do not impose this screen, which removes 80 funds.

(in June and December) and hence must be dropped if we are to proxy for trades based on quarterly changes in holdings.

Finally, we eliminate funds that do not have the same management company as the plurality of other funds in the same family. We do so because we infer that the management of these funds is contracted out, and hence may not take place in the same city as the family's headquarters. For example, one of the funds in the Vanguard family is the Wellington Fund, which is managed by the Wellington Management Co., whereas most Vanguard funds are managed by the Vanguard Group. While the Vanguard family is headquartered in Philadelphia, the Wellington Management Co. is located in Boston. Because it is important for our purposes to know where fund managers actually are located, observations such as the Wellington Fund are a potentially problematic source of measurement error.[8]

After all of these screens, we are left with 1,635 funds that appear in our sample as of December 1998.[9] Table I gives basic information on the distribution across cities of these funds and their respective fund families. We rank cities by the number of fund families that are located in each. We also tabulate the number of individual funds in each city, as well as each city's share of the mutual fund market, stated relative to both the total number of funds in our sample and the dollar value of all fund assets. As can be seen, New York and Boston dominate the mutual fund landscape. New York has the most families, with 114, and its 444 funds represent 27.2% of the 1,635 funds in our sample as of December 1998. Boston has fewer families and funds (42 and 324, respectively), but a much larger share of total fund assets than New York, by a margin of 41.9% to 16.0%. This is due to the fact that some of the very largest mutual funds are located in Boston. Other major mutual fund cities include Philadelphia (34 families, 5.0% of funds, 4.2% of total assets) and Los Angeles (26 families, 5.0% of funds, 11.5% of total assets).

In our baseline analysis, we look for peer group effects within the 15 "big" cities listed in Table I, defined as those cities that have six or more mutual fund families: New York, Boston, Philadelphia, Chicago, San Francisco, Los Angeles, Minneapolis, Baltimore, Atlanta, Milwaukee, Houston, Miami, Tampa, Denver, and Kansas City. The idea of this six-family cutoff is that there needs to be several fund families in the same city in order for fund managers to have a meaningful chance to interact with other fund managers outside their own family. The rest of the mutual funds that are not located in one of these focal cities are lumped together in "Rest of Cities"; this category comprises 18.8% of the funds but only 6.5% of fund assets. We also perform sensitivity checks,

[8] We are grateful to several seminar participants for emphasizing this point to us. In an earlier draft of the paper, we did not have this last screen, which cuts down the number of funds in our sample by about 14%. Our results were qualitatively similar to those we report here, but—as might be expected from a measurement-error problem—significantly smaller in magnitude, as well as more erratic on a city-by-city basis.

[9] The exact number of funds in our sample varies from one quarter to the next. For purposes of comparison, note that Coval and Moskowitz (1999) are left with 1,258 funds after applying a similar selection procedure to 1994 data.

**Table I**

**Cities Ranked by Number of Fund Families, December 1998**

This table reports summary statistics on the distribution across cities of those funds and fund families that are in our sample as of December 1998. We rank cities by the number of fund families that are located in each. We also tabulate the number of individual funds in each city, as well as each city's share of the mutual fund market, stated relative to both the total number of funds in our sample and the dollar value of all fund assets. Starting with the CDA Spectrum database, we only retain those U.S.-based stock funds for whom we are able to find a family location in the Nelson's Directory, and that report holdings on a quarterly basis. We also discard index funds, as well as those funds that have a management company that differs from that of the plurality of other funds in the same family. These screens leave us with a total of 1,635 funds in December 1998.

| | City | Number of Fund Families | Number of Funds | % of All Funds | % of All Fund Assets |
|---|---|---|---|---|---|
| 1 | New York | 114 | 444 | 27.2 | 16.0 |
| 2 | Boston | 42 | 324 | 19.8 | 41.9 |
| 3 | Philadelphia | 34 | 81 | 5.0 | 4.2 |
| 4 | Chicago | 30 | 85 | 5.2 | 2.3 |
| 5 | San Francisco | 27 | 64 | 3.9 | 1.3 |
| 6 | Los Angeles | 26 | 81 | 5.0 | 11.5 |
| 7 | Minneapolis | 15 | 41 | 2.5 | 0.5 |
| 8 | Baltimore | 12 | 44 | 2.7 | 3.7 |
| 9 | Atlanta | 12 | 12 | 0.7 | 0.3 |
| 10 | Milwaukee | 11 | 28 | 1.7 | 1.1 |
| 11 | Houston | 11 | 33 | 2.0 | 3.4 |
| 12 | Miami | 8 | 20 | 1.2 | 0.2 |
| 13 | Tampa | 7 | 7 | 0.4 | 0.0 |
| 14 | Denver | 6 | 32 | 2.0 | 3.3 |
| 15 | Kansas City | 6 | 31 | 1.9 | 3.8 |
| 16 | Rest of Cities | 114 | 308 | 18.8 | 6.5 |

where we include either 10 or 25 focal cities in our analysis. As detailed below, these variations lead to very similar results.

Our estimation strategy proceeds as follows. First, we define $h^i_{j,k,l,t}$ as the fractional share of its portfolio that mutual fund $j$, in family $k$ and located in city $l$, holds in stock $i$ in quarter $t$. Note that $h^i_{j,k,l,t}$ is only defined for the 15 big cities. Thus, each quarter, we have an observation for every stock/big-city-fund pair in our sample. In particular, if we restrict ourselves to the top 2,000 stocks in the CRSP universe—as we do in our baseline analysis—and given that we have 1,327 funds (out of the total of 1,635) located in our big cities in December 1998, there will be 2,654,000 observations on $h^i_{j,k,l,t}$ for this quarter. Of course, many of these observations will be zeros, since most funds only hold a small fraction of the available stocks.

Next, we define $H^i_{c,t}$ as the equally weighted average across all funds in city $c$ of the share of funds' portfolios invested in stock $i$ in quarter $t$. In other words, $H^i_{c,t}$ measures how heavily invested in stock $i$ is the typical mutual fund in city $c$. Similarly, $H^i_{c,xk,t}$ is the equally weighted average of the share of funds' portfolios invested in stock $i$ in quarter $t$, but now the averaging is done across all funds in city $c$ except those in family $k$. Finally, $H^i_{R,t}$ is the equally weighted average across all remaining funds—that is, those funds not in one of our focal big cities—of the share of funds' portfolios invested in stock $i$ in quarter $t$.

Lastly, we need to define a couple of indicator variables. We denote by $I(l = c)$ an indicator that takes the value of one if city $l$ and city $c$ are the same, and 0 otherwise. Complementarily, we denote by $I(l \neq c)$ an indicator that takes the value of 1 if city $l$ and city $c$ are different. In other words, we have $I(l \neq c) = (1 - I(l = c))$.

We are now ready to write down our baseline specification for trading behavior. We pool the data from March 1997 through December 1998, and run the following ordinary least squares (OLS) regression:

$$\Delta h^i_{j,k,l,t} = \sum_c \alpha_c \{\Delta H^i_{c,xk,t} \cdot I(l = c)\} + \sum_c \beta_c \{\Delta H^i_{c,t} \cdot I(l \neq c)\} + \gamma \Delta H^i_{R,t} + \varepsilon^i_{j,k,l,t}. \tag{1}$$

One noteworthy detail is that in calculating $\Delta h^i_{j,k,l,t}$ for any quarter $t$, we use only prices from quarter $t-1$. That is, for any stock $i$, we look at the change in the number of shares held between $t-1$ and $t$, and multiply this change by the price at $t-1$ in order to get the dollar change in holdings. The idea is to set aside any comovements in measured trades across funds that are mechanically induced by price changes, rather than by their active buying and selling decisions.

This specification can be interpreted as follows. For each of our big cities of interest, we estimate two coefficients: an "own-city" effect given by the $\alpha_c$'s, and an "other-city" effect given by the $\beta_c$'s. For example, if Boston corresponds to $c = 2$, then $\alpha_2$ measures how the trades of a given Boston fund are influenced by the trades of other Boston funds not in the same family, and $\beta_2$ measures how the trades of funds located in cities other than Boston are influenced by the trades of all Boston funds. The $\gamma \Delta H^i_{R,t}$ term allows the trades of the remaining funds—those not located in one of our focal big cities—to also have an influence on the trades of any one of our big-city funds.

Our key hypothesis is that the own-city effects should be larger than the other-city effects, that is, that for any $c$, we should have $\alpha_c > \beta_c$. For example, we expect the trades of a given Boston-based fund to be more responsive to the trades of other Boston funds (not in the same family) than would be the trades of a fund located outside of Boston. To boil everything down to a single test statistic, we take the weighted average of the $\alpha_c$'s and compare it to the weighted average of the $\beta_c$'s, with the weighting done according to the number of funds in each city.[10]

We also consider a couple of richer dynamic specifications, where we allow a given fund's trades in any quarter $t$ to be influenced by the trades of funds in

[10] We have experimented with other weighting schemes, such as weighting by the dollar value of fund assets in each city. The results are very similar to those we report below.

its own and other cities in the prior quarters $t-1$ and $t-2$. For example, in the case in which we only allow for one lag, the specification becomes

$$\begin{aligned}\Delta h^i_{j,k,l,t} = &\sum_c \alpha_{c,0}\{\Delta H^i_{c,xk,t}\cdot I(l=c)\} + \sum_c \beta_{c,0}\{\Delta H^i_{c,t}\cdot I(l\neq c)\}\\ &+\gamma_0\Delta H^i_{R,t} + \sum_c \alpha_{c,1}\{\Delta H^i_{c,xk,t-1}\cdot I(l=c)\}\\ &+\sum_c \beta_{c,1}\{\Delta H^i_{c,t-1}\cdot I(l\neq c)\} + \gamma_1\Delta H^i_{R,t-1} + \varepsilon^i_{j,k,l,t},\end{aligned} \tag{2}$$

where $\alpha_{c,0}$ and $\beta_{c,0}$ now denote the contemporaneous own-city and other-city coefficients for city $c$, while $\alpha_{c,1}$ and $\beta_{c,1}$ denote the once-lagged coefficients. The specification in which we allow for two lags is set up analogously.

It is worth contrasting our baseline specification in (1) with a simpler and more intuitive alternative. Let $H^i_{M,t}$ denote the equally weighted average across all funds of the share of funds' portfolios invested in stock $i$ in quarter $t$. One can run a regression of the form

$$\Delta h^i_{j,k,l,t} = a\cdot\Delta H^i_{c,xk,t}\cdot I(l=c) + b\cdot\Delta H^i_{M,t} + \varepsilon^i_{j,k,l,t}. \tag{3}$$

According to the specification in (3), the degree to which a given fund deviates from typical mutual fund sector-wide behavior (i.e., from $\Delta H^i_{M,t}$) in its trading of stock $i$ is allowed to be a function of how other funds in the same city (but not the same family) trade stock $i$. In this setup, the key test would be to ask whether $a > 0$. While this approach sounds appealingly simple and straightforward, it suffers from two drawbacks that lead us to prefer the more elaborate specification in (1).

First, specification (3) imposes the restriction that the influence of any given city on other cities—that is, its weight in the market portfolio variable $H^i_{M,t}$—is proportional to the number of funds located there. It is not a priori obvious that this is the right assumption. For example, it might be that the influence of a city is determined instead by the dollar value of assets in funds located there, or by other factors. By allowing the regression to pick out a different other-city coefficient (that is, a different $\beta_c$) for each of our cities of interest, specification (1) is more flexible in this regard.[11]

Second, and more significant, specification (3) suffers from a double-counting problem that can become a source of bias to the extent that some of the biggest cities in our sample comprise a substantial fraction of the entire mutual fund universe. To see the nature of this problem, suppose that consistent with a peer effects story, all of the funds in New York together increase their holdings of a particular stock $i$, thereby generating positive values of $\Delta h^i_{j,k,NY,t}$ and $\Delta H^i_{NY,xk,t}$. Ideally, we would like this sort of variation in the data to translate into a positive $a$ coefficient in a regression such as (3). But, since New York

[11] A related advantage of (1) is that it allows us to see whether each individual city is contributing in the right way to the overall results. That is, we can check if $\alpha_c > \beta_c$ in every city, something which cannot be done with (3).

comprises a large fraction of the market portfolio, $\Delta H^i_{M,t}$ also increases significantly. So the realization of $\Delta h^i_{j,k,\mathrm{NY},t}$ is partially "explained" by the realization of $\Delta H^i_{M,t}$, thereby reducing the apparent importance of $\Delta H^i_{NY,xk,t}$, and hence shrinking our estimate of $a$ toward 0. Intuitively, since the New York portfolio (as represented by $H^i_{NY,xk,t}$) and the market portfolio (as represented by $H^i_{M,t}$) are mechanically somewhat similar, the distinct influence of a change in the New York portfolio is estimated with a downward bias. Specification (1) does not suffer from this double-counting problem since it never uses a single market portfolio variable $H^i_{M,t}$. Instead, changes in the New York portfolio are simply allowed to enter with different weights when explaining the trades of funds in New York and elsewhere.

In one of our robustness checks below, we experiment with using specification (3) in place of (1). As expected, the results are somewhat attenuated, in the sense that the estimate of the $a$ coefficient in (3) is smaller than the weighted average difference between the $\alpha_c$'s and the $\beta_c$'s in (1). But the difference is not large—on the order of 15% or so. Thus, while specification (1) remains our preferred approach, it does not lead us to very different conclusions than what we get with a much simpler alternative.

In spite of its advantages, specification (1) still suffers from another subtle bias that could lead us to understate the difference between the $\alpha_c$'s and the $\beta_c$'s. This bias arises when a single family represents a large share of the funds in a particular city. To see how it works, consider the case of Denver, where the Janus family controls 41% of all funds. Now suppose we take an observation of the left-hand-side variable $\Delta h^i_{j,k,l,t}$ for a Janus fund—that is, an observation for which $k =$ Janus and $l =$ Denver. When estimating the own-city $\alpha$ coefficient for Denver, $\Delta h^i_{j,k,l,t}$ will be matched up with $\Delta H^i_{c,xk,t}$, the change in the share of non-Janus Denver funds' portfolios invested in stock $i$. In contrast, when estimating the other-city $\beta$ coefficient for Denver, non-Denver observations will be matched up with $\Delta H^i_{c,t}$, the change in the share of all Denver funds' portfolios invested in stock $i$.

The problem is that the latter independent variable is simply more informative about what mutual funds in general are doing with respect to stock $i$ because in not excluding Janus funds, $\Delta H^i_{c,t}$ gives a more accurate picture of changes in the "Denver portfolio" than does $\Delta H^i_{c,xk,t}$. In other words, we would expect $\Delta H^i_{c,t}$ to do a better job of explaining the trades of funds in an arbitrary other-city (say Miami) than $\Delta H^i_{c,xk,t}$, where in this case $\Delta H^i_{c,xk,t}$ is the change in the Denver portfolio excluding Janus. Thus, we have more of a classical errors-in-variables problem when estimating the own-city $\alpha$ coefficient than when estimating the other-city $\beta$ coefficient, which biases downward our estimate of the differential between the two.

We take a crude stab at mitigating this bias in another of our robustness checks. We do so by discarding all observations corresponding to families that control more than 20% of the funds in their home city. In the context of the previous example, this means not using any observations from Janus funds on the left-hand side of the regression. At the same time, however, the holdings

**Table II**

**Own-City and Other-City Effects in Mutual Fund Trades, March 1997 to December 1998**

This table reports the results for OLS regressions corresponding to equations (1) and (2) of the text. The dependent variable is the change in the percentage holding of a stock by a fund. For the no-lag specification in (1), we report in Panel B all of the own-city and other-city coefficients (the 15 $\alpha_c$'s, and the 15 $\beta_c$'s, respectively), as well as the rest-of-market coefficient $\gamma$. We summarize the results in Panel A, showing the weighted average of the difference between the $\alpha_c$'s and the $\beta_c$'s, with the weighting done according to the number of funds in each city. We also report in a similar manner the results for the specification in (2), which adds one lag. Here we show not only all of the contemporaneous own-city and other-city coefficients (the 15 $\alpha_{c,0}$'s, and the 15 $\beta_{c,0}$'s, respectively) but also the corresponding lagged coefficients (the 15 $\alpha_{c,1}$'s, and the 15 $\beta_{c,1}$'s, respectively). Our sample includes the top 2,000 stocks in the CRSP universe, and funds in those 15 cities that are home to six or more fund families. For a fund to be in the sample, it must report holdings data in consecutive quarters. The regressions pool all observations over the period of March 1997 to December 1998. The standard errors are adjusted to allow for correlation of observations within a stock cell.

Panel A: Summary

| | No Lags | One Lag | |
|---|---|---|---|
| | Contemporaneous Effect | Contemporaneous Effect | One-Quarter Lagged Effect |
| Weighted average difference own-city vs. other-city effects | 0.1310<br>(0.0108) | 0.1325<br>(0.0115) | 0.0212<br>(0.0154) |

(*continued*)

of Janus funds are still included in both the $\Delta H^i_{c,t}$ and $\Delta H^i_{c,xk,t}$ variables corresponding to Denver. Not surprisingly, this produces higher estimates of the own-city/other-city differential. But again, the differences relative to our baseline specification are modest.

## II. Results

### *A. Baseline Results for Trades*

Table II presents a detailed overview of our results for mutual fund trades. As noted above, our baseline sample includes the top 2,000 stocks in the CRSP universe, and all of the funds in those 15 cities (shown in Table I) that are home to six or more fund families.[12] For both specifications (1) and (2), we display in Panel B all of the own-city and other-city coefficients (the individual $\alpha_c$'s and $\beta_c$'s, respectively), as well as the rest-of-market coefficient $\gamma$. We also summarize the results in Panel A by computing the weighted average of the difference between the $\alpha_c$'s and the $\beta_c$'s, with the weighting done according to the number of funds in each city. This weighted average differential is our key test statistic.

[12] More precisely, a stock is included for all quarters in 1997 to 1998 if it ranks in the top 2,000 of the CRSP universe as of December 1998.

**Table II**—*Continued*

| Panel B: Details | | | | | | |
|---|---|---|---|---|---|---|
| | No Lags | | One Lag | | | |
| | Contemporaneous Effect | | Contemporaneous Effect | | One-Quarter Lagged Effect | |
| | Own-City | Other-City | Own-City | Other-City | Own-City | Other-City |
| New York | 0.2842 | 0.1441 | 0.2984 | 0.1541 | 0.0293 | 0.0212 |
| | (0.0304) | (0.0212) | (0.0304) | (0.0219) | (0.0165) | (0.0103) |
| Boston | 0.2491 | 0.1127 | 0.2521 | 0.1114 | 0.0580 | −0.0070 |
| | (0.0316) | (0.0185) | (.0326) | (0.0187) | (0.0190) | (0.0168) |
| Philadelphia | 0.2165 | 0.0497 | 0.2371 | 0.0511 | 0.0320 | 0.0042 |
| | (0.0255) | (0.0057) | (0.0292) | (0.0063) | (0.0137) | (0.0055) |
| Chicago | 0.1492 | 0.0490 | 0.1416 | 0.0530 | −0.0094 | −0.0057 |
| | (0.0163) | (0.0067) | (0.0236) | (0.0081) | (0.0223) | (0.0053) |
| Los Angeles | 0.1069 | 0.0422 | 0.1076 | 0.0422 | −0.0067 | 0.0027 |
| | (0.0168) | (0.0058) | (0.0195) | (0.0072) | (0.0148) | (0.0046) |
| San Francisco | 0.1377 | 0.0388 | 0.1277 | 0.0359 | 0.0002 | 0.0035 |
| | (0.0187) | (0.0054) | (0.0243) | (0.0056) | (0.0190) | (0.0040) |
| Minneapolis | 0.1467 | 0.0303 | 0.1566 | 0.0303 | 0.0242 | −0.0020 |
| | (0.0246) | (0.0053) | (0.0255) | (0.0059) | (0.0162) | (0.0034) |
| Baltimore | 0.0925 | 0.0279 | 0.0820 | 0.0341 | 0.0050 | 0.0033 |
| | (0.0148) | (0.0062) | (0.0158) | (0.0076) | (0.0091) | (0.0047) |
| Atlanta | 0.3047 | 0.0106 | 0.2626 | 0.0103 | 0.0480 | 0.0019 |
| | (0.0174) | (0.0022) | (0.0181) | (0.0023) | (0.0133) | (0.0014) |
| Milwaukee | 0.0738 | 0.0272 | 0.0817 | 0.0292 | −0.0061 | −0.0007 |
| | (0.0152) | (0.0029) | (0.0162) | (0.0034) | (0.0128) | (0.0031) |
| Houston | 0.1076 | 0.0074 | 0.1032 | 0.0086 | 0.0317 | 0.0044 |
| | (0.0209) | (0.0028) | (0.0226) | (0.0031) | (0.0121) | (0.0030) |
| Miami | 0.2566 | 0.0083 | 0.2519 | 0.0078 | 0.0067 | 0.0070 |
| | (0.0284) | (0.0025) | (0.0316) | (0.0026) | (0.0128) | (0.0032) |
| Tampa | 0.2312 | 0.0007 | 0.2347 | 0.0008 | 0.0264 | −0.0029 |
| | (0.0241) | (0.0011) | (0.0262) | (0.0010) | (0.0112) | (0.0015) |
| Denver | 0.2614 | 0.0173 | 0.2659 | 0.0191 | −0.0317 | −0.0017 |
| | (0.0244) | (0.0026) | (0.0311) | (0.0035) | (0.0277) | (0.0027) |
| Kansas City | 0.1391 | 0.0170 | 0.1155 | 0.0169 | 0.0299 | 0.0028 |
| | (0.0273) | (0.0074) | (0.0429) | (0.0074) | (0.0170) | (0.0024) |
| Rest of Market | | 0.1391 | | 0.1311 | | −0.0098 |
| | | (0.0151) | | (0.0155) | | (0.0085) |

The numbers in the first set of columns correspond to the no-lags specification in (1). The weighted average differential between the own-city coefficients and the other-city coefficients is 0.1310, with a $t$-statistic of 12.13.[13] This differential implies that for a given fund $j$, if all the other funds in $j$'s city (but not in $j$'s family) increase their purchases of a particular stock by 1 percentage point of total assets, then fund $j$ can be expected to increase its purchases of that stock by 0.1310 percentage points more than a fund in another city.

[13] The standard errors are adjusted to allow for arbitrary correlation of observations within a given stock cell.

Turning to the city-by-city detail, we can see that the own-city effect exceeds the other-city effect in each of the 15 cities under consideration, with the differences being statistically significant at the 5% level in every case. (These $p$-values are not shown in the table.) Moreover, even if one does not trust our standard errors, the fact that the point estimates go the right way 15 out of 15 times is strong nonparametric evidence against the null hypothesis: An outcome of 15 heads has only a 0.00003 probability of occurring if a fair coin is flipped 15 times.

To get a sense of magnitudes, consider the following thought experiment. For each stock $i$, we have in every quarter 15 observations of the average fund holdings at the city level, $H^i_{c,t}$. We demean these $H^i_{c,t}$'s, allowing for a separate mean for each stock. With 2,000 stocks, this creates a pool of 30,000 residuals each quarter. In December 1998, the standard deviation of these residuals is 0.0525 percentage points.[14] Thus, if the other funds in $j$'s city have portfolios that overweight stock $i$ by two standard deviations, then fund $j$ can be expected to overweight stock $i$ by 0.014 percentage points more than would a fund in another city ($0.0525 \times 2 \times 0.1310 = 0.014$). As a benchmark, the mean fraction of its portfolio that a fund holds in any one stock is 0.035%.[15] So a two-standard-deviation shock to city-level holdings has an incremental impact on fund $j$'s holdings of stock $i$ that is about 40% of its unconditional mean (0.014/0.035 = 0.40).[16]

This effect can be compared to the local-bias effect among fund managers that is documented by Coval and Moskowitz (1999). In our holdings regressions (discussed below), we find that when a fund and a stock are located in the same Census region, the fund increases its holdings of that stock by 0.004 percentage points of total assets, or by 11% of its unconditional mean holdings (0.004/0.035 = 0.11). Thus, the impact of a two-standard-deviation change in city-level holdings is roughly four times as large as the impact of a stock being local.

In the specification that adds one lag, the contemporaneous differential is roughly unchanged, at 0.1325, while the once-lagged differential (i.e., the weighted average value of $\alpha_{c,1} - \beta_{c,1}$) is much smaller, at 0.0212, and not statistically significant.[17] Thus, the quarterly data appear to be too coarse for us to say much about the details of the dynamic adjustment process. Clearly, it would be nice to have higher frequency data for this purpose.

[14] In calculating this standard deviation, we weight observations of city-level residuals by the number of funds in the city. Our estimates of economic magnitudes are actually somewhat larger if we give each residual equal weight.

[15] To put this figure in perspective, recall that most funds have zero holdings of most stocks. So the mean holding conditional on actually owning a given stock is much higher.

[16] Implicit in this experiment is the notion that the estimated differential of 0.1310 obtained from the changes specification in (1) can also be thought of as a conservative version of the estimate that would obtain in a levels version of the specification, as in equation (4) below. Table AI in the appendix verifies that this is indeed the case.

[17] A specification with two lags yields very similar results. In particular, the twice-lagged differential is again positive, but even smaller than the once-lagged differential.

**Table III**

**Own-City and Other-City Effects in Mutual Fund Trades: Quarterly Fama–MacBeth Regressions**

This table reports Fama–MacBeth results for OLS regressions corresponding to equation (1) of the text. The dependent variable is the change in the percentage holding of a stock by a fund. For each quarter, we display the weighted average of the difference between the $\alpha_c$'s and the $\beta_c$'s, with the weighting done according to the number of funds in each city. Our sample includes the top 2,000 stocks in the CRSP universe, and funds in those 15 cities that are home to six or more fund families. For a fund to be in the sample, it must report holdings data in consecutive quarters.

| | |
|---|---|
| 4th Quarter 1998 | 0.1376 |
| 3rd Quarter 1998 | 0.1742 |
| 2nd Quarter 1998 | 0.1085 |
| 1st Quarter 1998 | 0.1572 |
| 4th Quarter 1997 | 0.0886 |
| 3rd Quarter 1997 | 0.0639 |
| 2nd Quarter 1997 | 0.0911 |
| Mean of quarterly estimates | 0.1173 |
| (Standard error of mean) | (0.0152) |

### *B. A Fama–MacBeth Analysis of Trades*

In a further effort to ensure that some unknown error correlation structure is not leading us to overstate the precision of our results, Table III takes a Fama–MacBeth (1973) approach to the analysis of trades. Rather than pooling all the data from 1997 to 1998 together into a single regression, we now estimate the contemporaneous-changes specification in (1) over seven separate cross sections, which correspond to the quarters ending between June 1997 and December 1998. (For the sake of brevity, we omit much of the detail shown in Table II, and for each regression only display the key summary statistic, namely the weighted average differential between the own-city coefficients and the other-city coefficients.) Across the seven regressions, the mean value of the own-city/other-city differential is 0.1173, and the median is 0.1085. The highest value is 0.1742, and the lowest is 0.0639. Based on the dispersion of these point estimates, the mean of 0.1173 has a Fama–MacBeth standard error of 0.0152 and an associated $t$-statistic of 7.72. Thus, even with this conservative methodology, it appears that our basic results are strongly significant.

### *C. Different Samples and Specifications for Trades Regressions*

In Table IV, we rerun our basic no-lags trade regression, using a variety of alternative samples and specifications. (Again, we show only the key summary statistic, the weighted average differential between the own-city coefficients and the other-city coefficients.) Note that Row 1 of Table IV reproduces

**Table IV**

**Own-City and Other-City Effects in Mutual Fund Trades: Various Samples and Specifications**

This table reports results for OLS regressions corresponding to equation (1) of the text, except in Row 2, where we use the simple specification in equation (3), and in Row 5, where we run a probit for whether a fund buys a previously uncovered stock. The dependent variable (except in Row 5) is the change in the percentage holding of a stock by a fund. We report the weighted average of the difference between the $\alpha_c$'s and the $\beta_c$'s, with the weighting done according to the number of funds in each city. For a fund to be in the sample, it must have holdings in consecutive quarters. The regressions pool all observations over the period March 1997 to December 1998. The standard errors are adjusted to allow for correlation of observations within a stock cell. The composition of the sample varies across the rows of the table, as indicated.

| | |
|---|---|
| 1. Top 2,000 stocks, 15 cities | 0.1310 |
| | (0.0108) |
| 2. Simple specification | 0.1094 |
| | (0.0069) |
| 3. Excluding large families (>20% of a city) | 0.1355 |
| | (0.0126) |
| 4. Excluding local stocks | 0.1297 |
| | (0.0126) |
| 5. Extensive margin: probits for initiating holding | 1.331 |
| | (0.1211) |
| 6. Intensive margin only | 0.8454 |
| | (0.0842) |
| 7. Low book-to-market stocks | 0.1374 |
| | (0.0133) |
| 8. High book-to-market stocks | 0.1087 |
| | (0.0116) |
| 9. Top 1,000 stocks, 15 cities | 0.1442 |
| | (0.0118) |
| 10. Top 3,000 stocks, 15 cities | 0.1248 |
| | (0.0093) |
| 11. Micro-cap stocks (below top 3,000), 15 cities | 0.1045 |
| | (0.0171) |
| 12. Top 2,000 stocks, 10 cities | 0.1487 |
| | (0.0146) |
| 13. Top 2,000 stocks, 25 cities | 0.1285 |
| | (0.0100) |

the estimate from our baseline version of the regression in Table II for easy reference.

In Row 2, we switch to using the simpler specification in (3). Here we report the $a$ coefficient, which is directly analogous to the weighted average difference between the $\alpha_c$'s and the $\beta_c$'s in specification (1). As can be seen, there is a relatively small decline in the measured effect—the $a$ coefficient is 0.1094, as compared to the figure of 0.1310 that we get when estimating (1). Thus, it seems clear that our qualitative results are not dependent on the use of this more elaborate specification.

In Row 3, we revert to specification (1), but drop all the observations corresponding to those "large" families that control more than 20% of the funds in their home cities.[18] (We still use the holdings of these large families in the calculation of the $\Delta H^i_{c,t}$ and $\Delta H^i_{c,xk,t}$ variables.) As discussed in Section I, this is a crude way of mitigating a bias in our methodology, one that might otherwise lead us to understate the own-city/other-city differential. Consistent with this idea, the estimated differential rises a bit, from 0.1310 to 0.1355.

In Row 4, we keep everything else the same as in the baseline, but exclude from the sample all observations for which the fund and the stock are located in the same census region.[19] Note that this is more stringent than simply controlling for the possibility that fund managers might have a fixed preference for local stocks—indeed, by running everything in changes, all of our trades specifications have already differenced out any fixed local preference effect. In contrast, by completely throwing out local stocks, we can address the following sort of possibility: Suppose that fund managers in a given city have better access to local executives, who pass along private information about their companies. If fund managers act on this information, their trades in local stocks will tend to covary positively, even if on average they show no net preference for these stocks.[20]

Of course, the flip side is that by excluding local stocks, we may also be suppressing some of the evidence that is most naturally supportive of word-of-mouth effects. After all, it is possible that fund managers might spend a disproportionate fraction of their time talking with one another about the prospects for local companies, even apart from any contact they have with local executives. In any case, as can be seen in the table, the impact of this exclusion is almost negligible: The own-city/other-city differential drops from 0.1310 to 0.1297.

In Row 5, we use probit instead of OLS, and the left-hand-side variable is dichotomous, taking the value of 1 if, in quarter $t$, a fund initiates a position in a stock $i$ that it did not own in quarter $t-1$, and taking the value of 0 otherwise. If the fund already owned the stock in quarter $t-1$ (i.e., if $h^i_{j,k,l,t-1} > 0$), this observation is excluded from the sample for quarter $t$. Thus, we are asking to what extent our results hold when we look only at variation on the extensive margin—that is, a fund making a purchase that turns the fund from a nonowner of a stock into an owner—as opposed to looking at variation on both the extensive and intensive margins simultaneously.

[18] The excluded families, along with their city shares, are as follows. Boston: Fidelity, 26%; Los Angeles: Merrill Lynch, 20%; Minneapolis: First American, 23%; Baltimore: T. Rowe Price, 61%; Milwaukee: Strong Capital, 35%; Houston: AIM, 47%; Miami: Templeton, 37%, Mackenzie Financial, 29%; Tampa: Eagle Asset, 43%; Denver: Janus, 41%, Founders, 25%; Kansas City: American Century, 42%, Waddell & Reed, 36%.

[19] There are nine census regions: New England (CT, ME, NH, RI, VT); Middle Atlantic (NJ, NY, PA); East North Central (IL, IN, MI, OH, WI); West North Central (IA, KS, MN, MO, NE, ND, SD); South Atlantic (DE, FL, GA, MD, NC, SC, VA, WV); East South Central (AL, KY, MS, TN); West South Central (AR, LA, OK, TX); Mountain (AZ, CO, ID, MT, NE, NM, UT, WY); and Pacific (AK, CA, HI, OR, WA).

[20] Analogously, by excluding local stocks, we also address the possibility that fund managers are responding to stories in the local media (newspaper, TV, etc.) that focus disproportionately on hometown companies.

The results suggest that there is, in fact, a good deal of action on the extensive margin. The weighted average differential between the own-city coefficients and the other-city coefficients (now stated in terms of marginal effects) is 1.331, with a $t$-statistic of 10.99. We can assess the economic magnitude of this estimate in a manner analogous to above, using the fact that the standard deviation of $\Delta H^i_{c,t}$ is 0.02561 percentage points. This implies that if the other funds in $j$'s city are two standard deviations more aggressive in buying stock $i$ in quarter $t$, then fund $j$ has a probability of initiating coverage of stock $i$ that is 0.068% greater than a fund in another city ($0.02561 \times 2 \times 1.331 = 0.068$). The benchmark in this case is that the unconditional probability of a fund buying a previously uncovered stock in any given quarter is 0.420%. Thus, a two-standard-deviation shock to city-level interest has an incremental impact on fund $j$'s propensity to initiate coverage that is about 16% of its unconditional mean (0.068/0.420 = 0.16).

In Row 6, we do the complementary exercise of isolating variation on the intensive margin. To do so, we run exactly the same OLS regression as in Row 1, but now exclude any observations on $\Delta h^i_{j,k,l,t}$ for which $h^i_{j,k,l,t-1} = 0$. In other words, for any given stock, we only look at funds that already own the stock, and ask how they adjust their positions in response to the trades of other funds in the same city. In this much smaller subsample, the own-city/other-city differential is 0.8454 (with a $t$-statistic of 10.04)—several times larger than its unconditional value of 0.1310. This increase makes intuitive sense, given that most mutual funds tend to specialize in a well-defined subset of the investable universe. In particular, a fund that already owns a stock is likely to be more receptive to information about it than one that does not, simply because ownership suggests that the stock matches the fund's stated objectives (e.g., small-cap growth).

Overall, Rows 5 and 6 demonstrate that our initial results reflect significant contributions from both the extensive and intensive margins. That is, the choices of a fund's same-city neighbors appear to influence not only its decision as to when to first buy a given stock, but also its subsequent post-initiation trading in the stock.

In Rows 7 and 8, we run our baseline specification on two equal-sized subsamples, one representing glamour (low book-to-market) stocks, and the other representing value (high book-to-market stocks). Word-of-mouth communication could plausibly have a greater role to play among glamour stocks, whose valuations are likely to be more dependent on soft, intangible information. Thus, one might expect the own-city/other-city differential to be larger among glamour stocks. This turns out to be the case, by a nontrivial margin—the estimate is 0.1374 for glamour stocks, 26.4% above the estimate of 0.1087 for value stocks.

In Rows 9 and 10, we try changing the number of stocks under consideration. In Row 9, we restrict ourselves to the top 1,000 stocks in the CRSP universe, and in Row 10 we expand to the top 3,000.[21] In both cases, the estimates are

[21] The market caps of the 1,000$^{\text{th}}$, 2,000$^{\text{th}}$, and 3,000$^{\text{th}}$ stocks in December of 1998 are \$802.3 million, \$239.2 million, and \$97.9 million, respectively.

very close to those from our baseline sample. In Row 11, we look at an entirely new sample, composed of the remaining micro-cap stocks that are below the top 3,000 (as of December 1998, there were 2,003 stocks in this group, with a median capitalization of $33.8 million). A priori, it is not clear whether one should expect to find stronger or weaker effects among these tiny stocks. On the one hand, given that they tend to be less well-covered by analysts and the media, it might seem that word-of-mouth information transmission would have a greater role to play. On the other hand, we are looking for evidence of this word-of-mouth behavior among institutional investors, and institutions tend to avoid the smallest stocks.[22] As it turns out, the estimate for the micro-cap category is similar to that from the larger cap samples, with a value of 0.1045.[23]

In Rows 12 and 13, we go back to using the top 2,000 stocks, and instead vary the number of cities that we deem to be large enough for inclusion. In Row 12, we tighten the screen, focusing only on the top 10 cities listed in Table I; these cities all are home to at least 11 fund families. In Row 13, we loosen the screen, and allow into consideration the 25 cities with more than four families. Again, the results are close to those from our 15-city base case.

### *D. Results for Holdings*

In addition to the trades regressions, we also experiment extensively with various regressions for holdings. Our most basic specification for holdings is nothing more than a levels version of equation (1), namely,

$$h^i_{j,k,l} = \sum_c \alpha_c \{H^i_{c,xk} \cdot I(l = c)\} + \sum_c \beta_c \{H^i_c \cdot I(l \neq c)\} + \gamma H^i_R + \delta \cdot \mathrm{LOCAL}_{i,l} + \varepsilon^i_{j,k,l}. \tag{4}$$

The one noteworthy addition here is the dummy $\mathrm{LOCAL}_{i,l}$, which takes the value of 1 if stock $i$ is headquartered in the same census region as city $l$, and the value of 0 otherwise. This dummy is a control for the local-bias effect of Coval and Moskowitz (1999); it is effectively differenced out in our trades specifications, but needs to be inserted when working with holdings.

We also try three variations on (4), which add further controls. (Again, all of these controls are differenced out in the trades regressions.) In the first of these, the $\mathrm{LOCAL}_{i,l}$ term is replaced with 135 fund-city × firm-census-region dummies. In the second variation, we leave in the $\mathrm{LOCAL}_{i,l}$ term, but add 225 fund-city × investment-style dummies. In the third variation, we replace

[22] Using NYSE breakpoints, Gompers and Metrick (2001) report that the fraction of shares owned by institutions at year-end 1996 was 0.55, 0.50, 0.40, 0.34, and 0.13 for stocks in the largest through smallest market-cap quintiles, respectively. Most of the stocks below the top 3,000 come from the smallest size quintile.

[23] A direct comparison of the point estimates can be misleading, as there is much less variance in the right-hand-side variable among the micro-cap stocks—the December 1998 standard deviation of the demeaned $H^i_{c,t}$'s is now only 0.0092 percentage points, as opposed to its value of 0.0525 percentage points among the top 2,000 stocks. Nevertheless, the implied economic magnitude is still large in relative terms among micro-caps: A two-standard deviation shock to city-level holdings has an impact on fund $j$'s holdings of stock $i$ that is about 330% of its unconditional mean. This is because unconditionally, funds have very small holdings of micro-cap stocks.

the fund-city $\times$ investment-style dummies with 1,245 fund-city $\times$ industry dummies.

The general picture that emerges from all of these holdings regressions is very similar to that obtained from the trades regressions. Thus, to save space, we do not discuss the holdings results any further here. The interested reader is referred to Table AI in Appendix A, which provides a compact summary of many of these results, using a format analogous to that of Table IV. Our earlier NBER working paper (Hong, Kubik, and Stein (2003)) contains a much more detailed account of this material.

## III. Alternative Interpretations

Our analysis has been motivated by the hypothesis that mutual fund managers located in the same city engage in direct word-of-mouth communication with one another, sharing information and ideas about the stocks in which they invest. While the results thus far would appear to be consistent with this hypothesis, they also admit a couple of other alternative interpretations. We now discuss these alternative stories, and, where applicable, describe our efforts to discriminate amongst them.

### *A. The LIR Hypothesis*

As mentioned in the introduction, one alternative possibility is that fund managers obtain inside information directly from the executives of the companies in which they invest, and that this privileged information distribution happens at the city level, either through conferences, road shows, or other forms of local-investor-relations activity. Although it is hard to completely rule out this LIR effect—and given its surface plausibility, it is not clear one should—it is possible to cobble together several pieces of evidence which, taken collectively, suggest that the LIR effect is not the whole story behind our results.

Broadly speaking, our empirical strategy for discriminating against the LIR hypothesis is to look for alternative samples where we have good a priori reasons to think that investor relations activity is likely to play a lesser role than in our baseline sample. If our results in these alternative samples are close to those in our baseline sample, we can have some confidence that these results are not entirely the product of an LIR effect.

We have already touched on two such experiments above. First, recall that we try to completely exclude all local stocks from our regressions. This variation is relevant in light of the LIR hypothesis because to the extent that company executives pay visits to any fund managers at all, one would think that they would be especially inclined to visit nearby fund managers. In other words, investor relations activity should be most intense in a company's own backyard. Yet our estimates are hardly changed when we exclude local stocks.

Second, we find that the basic results are robust among smaller stocks. Arguably, smaller companies should be expected to devote less in the way of resources to the sort of institutionally oriented investor relations activity that we have in mind here. For one thing, the fixed costs of such activity ought to loom larger to these firms. Moreover, small-cap—and especially micro-cap—companies have shareholder bases that are significantly less institutional on average, and more heavily tilted toward individual investors.[24]

A third way to address the LIR hypothesis is to make use of a recent regulatory change. As noted in the introduction, Reg FD, which was approved by the SEC on August 10, 2000, and became effective as of October 23, 2000, is precisely intended to reduce the extent to which company executives can selectively communicate information to a subset of their investors. Reg FD specifically requires a company to reveal any material information to all investors and Wall Street analysts simultaneously in the case of intentional disclosures, and within 24 hours in the case of unintentional disclosures.

There is some evidence that Reg FD has accomplished its objectives. In a recent survey by the Association for Investment Management Research (2001), about 90% of sell-side analysts reported holding individual interviews with top company managers prior to Reg FD, but roughly 70% of these analysts reported a drop in such contact post–Reg FD. Other surveys reported in Agarwal and Chadha (2004) and Zitzewitz (2002) indicate that analysts and other institutional investors believe that their forecasts and recommendations to clients have been substantially impaired by Reg FD, and that Reg FD has led chief financial officers to reduce both the quality and quantity of their investor communications efforts.

In the spirit of an event study, we would like to verify if our basic results hold in the immediate aftermath of Reg FD. Thus, we rerun the contemporaneous version of the trades regression given in equation (1) over two further intervals: (i) September 2000 to December 2000, and (ii) December 2000 to March 2001.[25] These two regressions are directly comparable to those reported in the Fama–MacBeth analysis of Table III.

Interestingly, we obtain fairly similar results: The own-city/other-city differential is 0.0782 (standard error of 0.0171) for the fourth quarter of 2000, and 0.0820 (standard error of 0.0195) for the first quarter of 2001. Recall that across the seven regressions in Table III, we find a mean differential of 0.1173, and a median of 0.1085. Thus, while the post–Reg FD estimates are smaller than the pre–Reg FD mean and median, the differences are neither economically large

[24] As noted above, most of the stocks below the top 3,000 come from the smallest quintile based on NYSE breakpoints. Gompers and Metrick (2001) report that the fraction of shares owned by institutions in this quintile is 0.13 as of year-end 1996. This contrasts with a figure of 0.55 for the top market-cap quintile. Relatedly, 47% of firms in our sub-3,000 category have no analyst coverage at all. Even in the 2,000-to-3,000 class, the median number of analysts is only two, and 25% of firms have no coverage.

[25] One might take issue with the first of these two intervals, since Reg FD only became effective in the middle of the fourth quarter of 2000. However, Zitzewitz (2002) argues that Reg FD actually had its maximal impact in this quarter, in part because of the perception that it would be more vigorously enforced during the remaining tenure of its chief proponent, Arthur Levitt, at the SEC. Note too that as we move further away from our baseline sample period of December 1998, the coefficients become less directly comparable across regressions. One problem is that our data on funds' locations is as of December 1998. Thus, we have to discard all those funds that come into existence after this date, which implies that we lose an increasing fraction of the fund universe the farther forward we go in time.

nor statistically significant. If one believes that Reg FD has had a real impact on what executives can selectively disclose to fund managers—and the survey evidence suggests that it has—this finding would seem to suggest that the LIR effect is not of first-order importance.

Finally, another relevant piece of evidence comes from outside the mutual fund sector. In recent work, Gamble (2003) replicates our basic findings for holdings using data on individual investors. This data, which comes from the records of a large brokerage firm, includes not only investors' account information, but also their zip codes. Gamble creates "cities" by aggregating up to the three-digit zip code level, and then—for those cities in which there are a sufficient number of observations—runs a cross-sectional regression which is essentially identical to our specification (4). As we do, he finds a significantly positive own-city/other-city differential. Since executives are much less likely to travel around visiting individual investors, this again would appear to be inconsistent with the LIR effect being the dominant force driving our results.

### *B. Reputational Herding at the City Level*

Another alternative explanation for our results begins with a career-concerns-type herding model of the sort developed by Scharfstein and Stein (1990). In this setting, one mutual fund manager might mimic the actions of another not because she learns anything useful from talking with him, but rather because she is afraid that doing something different might be harmful to her labor market reputation. To date, all the empirical applications of this theory to the mutual fund industry (e.g., Chevalier and Ellison (1999)) have construed the labor market in question to be a national one. In other words, the empirical work has implicitly assumed that any given fund manager has an equal incentive to mimic those in cities close to herself, and those further away. But one can imagine reformulating the theory slightly so that there is a greater incentive to mimic those more proximate. If fund managers tend to have limited mobility across cities, then a Boston-based manager's wage may be more closely linked to her relative standing in the Boston community than to her relative standing in the national community.

To put the distinction with the word-of-mouth model most sharply, observe that this version of the reputational herding mechanism can in principle apply even if all fund managers across the country have access to the same information. A Boston-based manager can know just as much about what Los Angeles managers are doing (and why they are doing it) as she knows about other Boston managers, but she will still tend to follow the Boston managers more closely because they are the ones that she is being benchmarked against.

However, it is also possible to construct stories that combine elements of both the reputational herding and word-of-mouth theories. For example, it might be that the labor market is a national one, so that for reputational reasons, a given manager only cares about her performance relative to those of all others in the industry. But at the same time, she might find it easier to get prompt information about what nearby managers are doing, so in her efforts to mimic current industry-wide holdings, she ends up with a portfolio that is especially close to those of managers who are located in the same city.

These sorts of hybrid stories suggest that it will be difficult to construct tests that can sharply discriminate between reputational herding and word-of-mouth effects in the mutual fund data. Following, for example, Chevalier and Ellison (1999), one might try to make some progress by adopting the identifying assumption that the incentives for reputational herding are greatest among younger money managers. But suppose that one then finds that younger managers are indeed the most prone to mimic the holdings and trades of their same-city counterparts. While this would be consistent with reputational herding playing some role in our results, it would not in any way rule out the importance of word-of-mouth communication for the reasons outlined above. That is, it may be that reputational herding among young managers plays itself out at the city—rather than national—level precisely because word-of-mouth information transmission makes it easier to implement a strategy of copying those who are close.

Again, however, if one is willing to look to data from outside the mutual fund sector, Gamble's (2003) study of individual investor behavior can be helpful in distinguishing among the stories. Since career concerns are presumably not an issue for individual investors, the fact that Gamble obtains results that so closely parallel our own makes it harder to argue for an interpretation of our evidence based solely on career concerns.

## IV. Conclusions

Our basic findings are easily summarized. The trades of any given fund manager respond more sensitively to the trades of other managers in the same city than to the trades of managers in other cities. This regularity is distinct from a local preference effect—indeed, it emerges even when local stocks are completely excluded from the analysis. The own-city/other-city differential is also pronounced across a wide range of market-cap categories.

This paper can be seen as part of a small recent literature that examines how word-of-mouth effects influence behavior in various financial market settings, joining, for example, Duflo and Saez (2002), Madrian and Shea (2000), Kelly and Grada (2000), and Hong, Kubik, and Stein (2004). However, one important distinction is that we do not know of any previous work that has posed the word-of-mouth question in the context of professional money managers. In the spirit of our opening quote from Shiller (2000), one of the main reasons to be interested in the word-of-mouth phenomenon in the first place is the possibility that it might ultimately have nontrivial implications for stock prices.[26] But if one is

[26] For example, word-of-mouth effects may turn out to be helpful in thinking about medium-horizon momentum in stock returns (Jegadeesh and Titman (1993)). According to Barberis, Shleifer, and Vishny (1998) and Daniel, Hirshleifer, and Subrahmanyam (1998), momentum arises when the representative investor misreacts to a piece of information. In contrast, Hong and Stein (1999) and Hong, Lim, and Stein (2000) argue that momentum comes from information that diffuses gradually across a heterogeneous set of investors (see also Chan (2003)). Our results would seem to be suggestive of a process of gradual information diffusion, though they do not speak to the momentum phenomenon directly.

only ever able to find evidence of word-of-mouth behavior among individuals, there will always be the objection that individuals are by themselves unlikely to exert a significant influence on stock prices, given the potentially powerful offsetting effects of professional arbitrage. Of course, our findings regarding trading behavior do not by themselves establish a link between word-of-mouth communication and stock prices. The only hope is that they leave the door open a bit wider than it was before, and thereby encourage further work on this topic.

## Appendix A

**Table AI**

**Own-City and Other-City Effects in Mutual Fund Holdings: Various Samples and Specifications**

This table reports results for OLS regressions corresponding to equation (4) of the text, except in Row 2, where we use a simpler specification that is an analog to equation (2), and in Row 5, where we run a probit for whether a fund holds a stock or not. The dependent variable (except in Row 5) is the percentage holding of a stock by a fund. Along with the baseline version of each regression, there are also three variations that include geographic controls, fund-style controls, and industry controls, respectively. The geographic controls are 15 fund-city dummies interacted with 9 firm-census-region dummies. The style controls are 15 fund-city dummies interacted with a set of 15 style dummies for each stock, which capture its location in size/book-to-market/return-momentum space. The industry controls are 15 fund-city dummies interacted with a set of 83 industry dummies (based on two-digit SIC codes). We report the weighted average of the difference between the $\alpha_c$'s and the $\beta_c$'s, with the weighting done according to the number of funds in each city. The standard errors are adjusted to allow for correlation of observations within a stock cell. The composition of the sample varies across the rows of the table, as indicated.

| | Baseline Model | Geographic Controls | Style Controls | Industry Controls |
|---|---|---|---|---|
| 1. Top 2,000 stocks, 15 cities, December 1998 | 0.2719 | 0.2676 | 0.2593 | 0.2401 |
| | (0.0158) | (0.0173) | (0.0185) | (0.0181) |
| 2. Simple specification, December 1998 | 0.2347 | 0.2275 | 0.2128 | 0.1970 |
| | (0.0219) | (0.0221) | (0.0226) | (0.0208) |
| 3. Excluding large families (>20% of a city), December 1998 | 0.2821 | 0.2768 | 0.2704 | 0.2458 |
| | (0.0240) | (0.0266) | (0.0280) | (0.0265) |
| 4. Excluding local stocks, December 1998 | 0.2644 | 0.2596 | 0.2490 | 0.2299 |
| | (0.0196) | (0.0213) | (0.0228) | (0.0223) |
| 5. Probits for holding, December 1998 | 3.124 | 2.649 | 1.899 | 1.747 |
| | (0.3139) | (0.3119) | (0.3295) | (0.3210) |
| 6. Low book-to-market stocks | 0.2770 | 0.2718 | 0.2621 | 0.2234 |
| | (0.0184) | (0.0202) | (0.0207) | (0.0216) |
| 7. High book-to-market stocks | 0.2110 | 0.1983 | 0.1872 | 0.1378 |
| | (0.0200) | (0.0212) | (0.0188) | (0.0229) |
| 8. Top 1,000 stocks, 15 cities, December 1998 | 0.2732 | 0.2644 | 0.2541 | 0.2146 |
| | (0.0165) | (0.0189) | (0.0192) | (0.0199) |
| 9. Top 3,000 stocks, 15 cities, December 1998 | 0.2683 | 0.2645 | 0.2567 | 0.2353 |
| | (0.0140) | (0.0162) | (0.0180) | (0.0174) |
| 10. Micro-caps (below top 3,000), 15 cities, December 1998 | 0.1932 | 0.1888 | 0.1899 | 0.1366 |
| | (0.0192) | (0.0194) | (0.0193) | (0.0221) |

(*continued*)

**Table AI**—*Continued*

| | Baseline Model | Geographic Controls | Style Controls | Industry Controls |
|---|---|---|---|---|
| 11. Top 2,000 stocks, 10 cities, December 1998 | 0.2789 | 0.2755 | 0.2673 | 0.2490 |
| | (0.0176) | (0.0194) | (0.0197) | (0.0191) |
| 12. Top 2,000 stocks, 25 cities, December 1998 | 0.2622 | 0.2604 | 0.2518 | 0.2389 |
| | (.0147) | (0.0167) | (0.0173) | (0.0169) |
| 13. Top 2,000 stocks, 15 cities, June 1998 | 0.2496 | 0.2445 | 0.2354 | 0.2204 |
| | (0.0149) | (0.0164) | (0.0175) | (0.0170) |
| 14. Top 2,000 stocks, 15 cities, December 1997 | 0.2489 | 0.2441 | 0.2412 | 0.2145 |
| | (0.0181) | (0.0201) | (0.0220) | (0.0204) |
| 15. Top 2,000 stocks, 15 cities, June 1997 | 0.2066 | 0.2010 | 0.1912 | 0.1689 |
| | (0.0184) | (0.0203) | (0.0225) | (0.0207) |

## REFERENCES

Agarwal, Anup, and Sahiba Chadha, 2004, Who is afraid of Reg FD? The behavior and performance of sell-side analysts following the SEC's Fair Disclosure rules, University of Alabama working paper.

Association for Investment Management Research, 2001, Regulation FD e-survey summary, available at http://www.aimr.com/pressroom/01releases/regfd_surveysum.htm.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein, 2005, Does function follow organizational form? Evidence from the lending practices of large and small banks, *Journal of Financial Economics* 76, 237–269.

Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan, 2000, Network effects and welfare cultures, *Quarterly Journal of Economics* 115, 1019–1055.

Case, Anne, and Lawrence Katz, 1991, The company you keep: The effect of family and neighborhood on disadvantaged youths, NBER working paper.

Chan, Wesley S., 2003, Stock price reaction to news and no-news: Drift and reversal after headlines, *Journal of Financial Economics* 70, 223–260.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith A., and Glenn D. Ellison, 1999, Career concerns of mutual fund managers, *Quarterly Journal of Economics* 114, 389–432.

Christoffersen, Susan, and Sergei Sarkissian, 2002, Location overconfidence, McGill University working paper.

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

Coval, Joshua D., and Tobias J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 4, 811–841.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A theory of overconfidence self-attribution and security market under- and over-reactions, *Journal of Finance* 53, 1839–1885.

Duflo, Esther, and Emmanuel Saez, 2002, Participation and investment decisions in a retirement plan: The influence of colleagues' choices, *Journal of Public Economics* 85, 121–148.

Dumais, Guy, Glenn D. Ellison, and Edward L. Glaeser, 2002, Geographic concentration as a dynamic process, *Review of Economics and Statistics* 84, 193–204.

Ellison, Glenn D., and Drew Fudenberg, 1995, Word of mouth communication and social learning, *Quarterly Journal of Economics* 110, 93–125.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gamble, Keith D., 2003, Word-of-mouth influence on the stock holdings of individual investors, Undergraduate thesis, Harvard University.

Glaeser, Edward L., Bruce Sacerdote, and Jose Scheinkman, 1996, Crime and social interactions, *Quarterly Journal of Economics* 111, 502–548.

Glaeser, Edward L., and Jose Scheinkman, 2002, Non-market interactions, in M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds.: *Advances in Economics and Econometrics: Theory and Applications*, Eighth World Congress (Cambridge University Press, Cambridge UK).

Gompers, Paul A., and Andrew Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229–259.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, 2000, Security analysts' career concerns and the herding of earnings forecasts, *RAND Journal of Economics* 31, 121–144.

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, NBER working paper 9711.

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

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 93–130.

Kelly, Morgan, and Cormac O. Grada, 2000, Market contagion: Evidence from the panics of 1854 and 1857, *American Economic Review* 90, 1110–1124.

Madrian, Brigette, and Dennis Shea, 2000, Peer effects and savings behavior in employer sponsored savings plans, University of Chicago working paper.

Petersen, Mitchell A., and Raghuram G. Rajan, 2002, Does distance still matter? The information revolution in small business lending, *Journal of Finance* 57, 2507–2532.

Rosenthal, Stuart S., and William C. Strange, 2001, The determinants of agglomeration, *Journal of Urban Economics* 50, 191–229.

Scharfstein, David S., and Jeremy C. Stein, 1990, Herd behavior and investment, *American Economic Review* 80, 465–479.

Shiller, Robert J., 2000, *Irrational Exuberance* (Broadway Books, New York).

Shiller, Robert J., and John Pound, 1989, Survey evidence on diffusion of interest and information among investors, *Journal of Economic Behavior and Organization* 12, 47–66.

Zitzewitz, Eric, 2002, Regulation Fair Disclosure and the private information of analysts, Stanford University working paper.