The Geography of S&P 500 Stock Returns*

David Barker and Tim Loughran†

February 28, 2007

Forthcoming, Journal of Behavioral Finance

Abstract

Investor bias in favor of geographically close firms has been documented in previous papers. An implication of this bias is that, if local events cause nearby investors to trade together, then the correlation of stock returns of pairs of firms will increase as the distance between them decreases. We test this hypothesis using a sample of Standard & Poors 500 (S&P) companies. After adjusting for industry effects and other factors, we find that the correlation coefficient between two stocks increases by 12 basis points for every 100 mile reduction in distance. This result is consistent with local shocks affecting the returns of nearby firms by an average of approximately 43 basis points per month. We conclude that these shocks are most likely the result of trading activity by local investors who own shares in nearby firms.

*We are grateful to Sebastian Rubano for research assistance and to David Bates, Gene Savin and Paul Schultz for their comments and suggestions. We also thank seminar participants at the Mendoza College of Business, University of Notre Dame and the Tippie College of Business, University of Iowa.

†**Contact Information:** Barker, dbarker@uiowa.edu, Tippie College of Business, University of Iowa, Iowa City, IA 52242 and Loughran, Loughran.9@nd.edu, Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556-5646.

Several recent papers have presented evidence that investment decisions of retail and institutional investors are influenced by location. In other words, individuals and institutions invest more than would otherwise be expected in firms that are geographically close to them. An implication of this finding is that the correlation of stock returns declines with distance between firms. In this paper, we examine the relationship between firm location and monthly return correlations of Standard and Poor's 500 Index (S&P 500) stocks during 2000-2004.

We find that geographic proximity plays an important role in monthly return correlations. Even after controlling for industry, differing analyst coverage, market capitalization, and other factors, as the geographic distance between two S&P 500 firms increases, the return correlation between the companies drops. For every 100 mile reduction in distance between firms, our results suggest that the correlation coefficient between the returns of the two stocks increases by 12 basis points, consistent with local shocks affecting the returns of nearby firms by an average of approximately 43 basis points per month. Correlation declines with distance as far as 2,000 miles, but beyond this distance the relationship appears to be different. As in much of the previous literature, we define firm location as the site of the firm's corporate headquarters. For example, we define Walt Disney Inc.'s firm location as Burbank, California.

Previous empirical evidence has shown that people tend to invest money in firms that they are familiar with, which are often companies located geographically close to them. Local media coverage, word-of-mouth information from neighbors or friends, or simply the ability to drive by the firm's headquarters appears to cause this strong bias in investment decisions. This behavioral bias should be minimized in large firms with wide geographic scope.

For this reason, we have selected the firms of the S&P 500 as a sample with which to test our hypothesis. The typical S&P 500 firm is large capitalization, well-followed by analysts, held primarily by institutions, and national or international in scope. These firms would seem to be the least likely to be affected by local events or the biases of local investors. By design, the stocks within the index span numerous different industries, from Pharmaceuticals to Real Estate. In addition, the firms of the index are diversified geographically, with firm headquarters located in more than 37 different states. If location can be shown to affect return correlations for these large market capitalization firms, then the local bias of investors would appear to be an important and widespread phenomenon.

But how would bias in favor of local firms affect the correlation of stock returns? If investors prefer familiar local companies, then firms that are close to each other are likely to have more investors in common than distant firms. Local economic and other considerations may disproportionately affect investors who live close to the same S&P 500 firms. These items may be local tax issues, news about the health of the local economy, or even weather, and they might produce local informational advantages, changes in local sentiment, or liquidity needs. If these factors cause local investors to trade together and if this common trading activity is large enough to affect prices, then stock returns of geographically close companies are likely to be correlated.

There are other reasons that return correlation might decline with distance. One is that similar firms tend to cluster geographically. If clustering is by industry, then nearby firms will be correlated simply because they are more likely to be in the same industry. We deal with this possibility by controlling for industry and other firm characteristics. Return correlation might also decline with distance if local economic conditions affect the fortunes of all nearby companies. We believe that our use of S&P 500 firms for our sample minimizes the chances that local conditions are driving the correlations of stock returns.

These results are important because they demonstrate that geographic location has a large effect on stock returns of major U.S. companies. Another paper, Pirinsky and Wang (2006), also found evidence that location matters for stock returns, but it contained serious methodological problems, and did not show that location affects the returns of large firms.

Our paper differs in several respects from Pirinsky and Wang (2006). Instead of directly measuring the effect of distance between firms as we do, Pirinsky and Wang (2006) group firms by MSA and regress stock returns on an index of other firms in the MSA, the market return, the return of a broad industrial group to which the firm belongs, and the returns of the two other industrial groups that are most closely correlated to the firm's returns. This regression is run for each firm in their sample, providing them with estimated coefficients (betas) for each firm on the index of local returns. They then compute the mean and standard deviation of the estimated coefficients across all of the firms in their sample. The statistical significance of their results, measured as the ratio of the mean to the standard deviation of the estimated coefficients, is enormous; the average t-statistic on the local index is 23.9.

The strength of the results of Pirinsky and Wang (2006) is probably due in part to two factors; non-independent observations and sub-industry correlation. It is likely that these factors have biased the mean of the estimated coefficients of their regressions upward and the standard deviation downward.

The observations that lack independence are the estimated coefficients of the firm-by-firm regressions. Suppose, for example, that there are only two firms in an MSA; firm A and firm B. Firm returns are regressed on an index of returns of local firms, so one regression will be run with firm A returns as the dependent variable and firm B returns as the independent variable, and another will be run with firm B returns as the dependent variable and firm A returns as the independent variable. The coefficients of these two regressions will clearly be correlated by construction, but they are reported in Pirinsky and Wang (2006) as independent trials. A t-statistic calculated as the ratio of the mean to the standard deviation of the estimated coefficients is only valid if the estimated coefficients are independent of each other. A large number of firms per MSA would reduce the problem, but the median number of firms per MSA in the Pirinsky and Wang (2006) sample is only 18. The minimum number is 5. Our methodology also presents a problem of non-independent observations, which we correct by using bootstrap regressions.

The industrial groups used by Pirinsky and Wang (2006) are from Fama and French (1997), which defines 48 broad industries. We find that there is significant geographic clustering by sub-industry within these groups. Failure to control for smaller industrial group correlation under these circumstances will magnify the reported effect of location. Pirinsky and Wang (2006) report that their results are unchanged using 2-digit SIC industry definitions, but the 8-digit GICS industry definitions we use are more detailed.

The potential for magnification of statistical significance in the methodology of Pirinsky and Wang (2006) is illustrated by their results on the comovement of firm earnings.[1] They find that changes in firm earnings are negatively correlated with an index of changes in earnings of local firms with t-statistics ranging from -23.4 to -29.1. Because the results are non-positive, they are reported as a "lack of local comovement in firm earnings." While is it very possible that they are correct that there is no relationship between firm earnings and earnings of nearby firms, the fact that Pirinsky and Wang (2006) find such a strong negative relationship indicates a problem with their methodology.

Another difference between our papers is the sample of firms used. Pirinsky and Wang (2006), use a large sample of 4,000 firms. The effect they find is strongest for small firms, and it is not clear from their results whether the effect exists at all for large firms. The stocks of the S&P 500, which we use, are all several times larger than the

[1]Reported in Table IV of Pirinsky and Wang (2006).

median firm in the sample of Pirinsky and Wang (2006).

The remainder of the paper is organized as follows. Section 1 provides a literature review. The data used here is described in Section 2. In Section 3 we provide the empirical results. We summarize our results and conclude in the last section.

1. Literature Review

A. Investor-Firm Proximity

The uncontroversial finding of the literature on investor proximity is that both mutual fund managers and retail investors disproportionately hold stock in their portfolio that is located geographically close to them. The more controversial result is that this bias in holding local stocks results in superior stock performance for investors. That is, the local bias may be driven by informational advantages, not just familiarity.

Using a unique Finnish dataset, Grinblatt and Keloharju (2001) report that the distance between firms and investors is an important factor in deciding which stocks are held by investors. Finnish investors (especially less sophisticated ones) exhibit a bias towards firms that are close, which publish annual reports in their native tongue, and have CEOs with similar cultural backgrounds as themselves. Institutional traders in Finland have less of a distance bias than individual investors.

Huberman (2001) finds that U.S. investors disproportionately own shares in the Regional Bell Operating Company (RBOC) that service their area at the diversification expense of holding shares of the other six RBOCs. He attributes this tendency to the fact that people prefer to invest in the familiar.

This preference for holding local stocks is also present in the trading habits of sophisticated U.S. money managers. For example, Coval and Moskowitz (1999) find that one of every 10 stocks held in a mutual fund portfolio is due to the firm's local location. In a follow-on paper, Coval and Moskowitz (2001) find that fund managers appear to be exploiting informational advantages by overweighting local stocks. The two authors find the average mutual fund manager earns 2.67% more per year on local stocks than on non-local holdings.

While Coval and Moskowitz (1999, 2001) document a preference for local companies by U.S. mutual fund managers, the local holding bias is even stronger for retail investors. Using a large sample of U.S. retail investors of a large discount broker, Ivkovic and Weisbenner (2005) report that the average share of local investments is 30%. They find that the average household earns about 3.2% more from its local holdings relative to non-local investments.[2] Interestingly, the authors find that no abnormal return is earned by the household's holding of local S&P firms.

Evidence consistent with word-of-mouth effects on the dissemination of information is presented in Hong, Kubik, and Stein (2003). They show that beyond local preference effects, U.S. mutual fund managers tend to be influenced by the buying and selling habits of other managers located in the same city. If one fund manager is increasing their position in a stock by 2.0%, managers from a different fund family in the same city will tend to increase their own holding in the stock by 0.4%.

Another paper that is suggestive of an effect of investor proximity to firms is Brown, Ivkovic, Smith, and Weisbenner (2004). These authors find that residents of communities that are near publicly traded firms are more likely to own stocks.

The only paper of which we are aware that investigates the effect of distance on correlation of stocks is Pirinsky and Wang (2006). As discussed in the introduction of this

[2] Zhu (2002) using a similar dataset, also finds strong evidence that retail investors overweight their portfolios in local stocks. Consistent with a familiarity explanation, he finds that advantageous information cannot explain the investor local bias.

paper, Pirinsky and Wang (2006) contains serious methodological problems, and does not show that location affects the returns of large firms. An interesting result of Pirinsky and Wang (2006), however, is that there does not appear to be any correlation of earnings of firms and local economic conditions. When indices of local and national economic fundamentals are added to their regressions, the t-statistic of the local economic index is -0.46. Given that the methodology of this paper tends to magnify statistical significance, this is convincing evidence that even small firms are not affected by economic conditions of the region surrounding their headquarters. If this is the case, then any relationship that is found between distance and correlation seems more likely to be the result of local trading activity than of the state of local economies.

In summary, the evidence is strong that investors disproportionately hold stocks that have captured their attention, which are often stocks that are located geographically close to them. An important logical consequence of local investor bias that has yet to be fully investigated is a relationship between stock return correlation and distance between firms.

B. Industrial Clustering

Another theory that might explain the relationship between stock returns and distance is industrial clustering. Theories of clustering date at least to Marshall (1920) who described the agglomeration economies that might lead similar firms to locate near each other. More recently a large literature has developed to describe and explain industrial clusters (Enright 2003, Porter 2000, Glenn and Glaeser, 1997). Firms cluster together to supply each other, share ideas, and share sources of labor and supplies. Some of these clusters are described as consisting of firms from different, but related industries. Clusters might also result from historical accidents, such as the home location of the founder of a firm, and the subsequent start-up of other firms by individuals leaving the original firm. Audretsch and Feldman (1996) find evidence that knowledge spillovers can create and be caused by geographic concentration, while Audretsch and Stephan (1996) stress that the knowledge spillovers that are most likely to be important for geographic concentration are those resulting from informal, unplanned interaction.

Previous literature gives us two theories that might explain why stock returns are correlated with distance; industry clustering and local investor bias. Properly controlling for industrial groups is clearly important if we are to distinguish between these two theories. In the analysis that follows, we control for industry effects by using the Global Industry Classification Standard (GICS) developed by Morgan Stanley Capital International and Standard & Poor's. Bhojraj, Lee, and Oler (2003) found the GICS definitions were better at explaining stock return comovements than other industry definitions.

The GICS defines broad and narrow industrial groups. Broad groups are given a 2-digit code; narrow groups are given an 8-digit code. Intermediate groups are given 4 and 6-digit codes. We use industry correlations at all levels to explain correlations between the returns of individual stocks, and then check to see if distance has any additional explanatory power. By using industry correlations at all levels we hope to capture the effects of the loose geographic groupings of companies in a variety of related industries described by Porter (2000) and isolate the effect of distance from industrial grouping.

2. Data

Although there are always 500 firms in the S&P 500 Index, our sample is restricted to those companies that are primarily based in the U.S. and that have a complete time series of monthly returns during 2000-2004. We use the University of Chicago's Center for Research in Security Prices (CRSP) for the sample monthly stock returns, shares outstanding, and share price information. Analyst coverage information is obtained from the Institutional Brokers' Estimates System (I/B/E/S). We define analyst coverage as the number of analysts reporting current fiscal year annual earning estimates each month during the sample period. Yahoo Finance serves as the source of the corporate headquarter addresses. Several other papers (see Coval and Moskowitz (1999), Zhu (2002), Ivkovic and Weisbenner (2005), Loughran and Schultz (2004, 2005)) also use the headquarters' location as the firm's location.

After determining the street addresses of the headquarters of the firms in our sample, we convert this information into latitude and longitude coordinates. Using these coordinates, we calculate the distance between every possible pair of firms. When calculating distances the size of the continental United States the curvature of the earth is a significant factor, so we use a formula that takes this into account. Determining the precise location of each firm allows us to investigate a wide range of distances, including firms that are located in the same building, across town, or across the country.

We use firms in the S&P 500 Index as of August 18, 2005 according to the S&P's website. Three firms are removed from the sample because of operational location in Bermuda (ACE Limited and XL Capital) or Montreal, Canada (Molson Coors Brewing Company).[3] In addition, several firms are removed because they did not have a complete series of 60 monthly stock returns according to CRSP.[4] In all of our analysis, the 456 firms in our final sample are equally weighted.

Panel A of Table 1 contains descriptive statistics on the data used in our sample of 456 firms. As noted earlier, S&P 500 firms have large market capitalizations (shares

[3]When firms are incorporated in Bermuda but have operational headquarters in the U.S. we use the U.S. location. For example, Tyco International is incorporated in Bermuda, but numerous press reports indicate that their operational headquarters are located in Princeton, New Jersey.

[4]There are a variety of reasons for firms to have at least one missing monthly return. Here are two examples. MetLife Inc. went public on April 4, 2000, so it is missing four monthly returns and is excluded from the sample. Rockwell Collins was spun off to the shareholders of Rockwell International in July of 2001. Hence, Rockwell Collins is missing 19 months of data and is also excluded from the final sample of firms.

outstanding multiplied by stock price), with a mean value of $19 billion and a median value of $8 billion. Typical insider ownership percentage is relatively low, which is not surprising given the sample's large market value. Institutional ownership is quite high, with institutions holding, on average, almost 75% of all shares outstanding. In addition, the average number of analysts covering the firms is approximately 16. Populations of metropolitan areas in which headquarters are located vary from 12,672 (Warsaw, Indiana, headquarters of Biomet, Inc.) to over 18 million (New York City).

Figure 1 shows the location of the S&P 500 firms that are in our sample. There is a large cluster of firms in the urban northeast and a vast area in the western portion of the country with very few firm headquarters. Almost all large metropolitan areas of the U.S. have at least one S&P 500 firm.

Panel B of Table 1 reports the sample characteristics for information technology firms while Panel C reports the characteristics for non-information technology firms. Information technology firms are defined as belonging to industry 45 of the Global Industry Classification Standard (GICS). This category includes software, services such as data processing and consulting, and manufacturing of computer and communications equipment. We examine information technology firms separately for two reasons. First, the location pattern of these firms appears to be different than that of firms in other industries in that information technology firms tend to cluster on both coasts. Second, information technology firms had an unusually high level of correlation during our sample period of 2000-2004, probably because of the collapse of the boom in technology stocks during these years.[5]

Figures 2 and 3 shows the location of information technology and non-information technology firms. Both types of firms are scattered across the country and are located in large and small cities, but information technology firms have a greater tendency to

[5] The Nasdaq Composite Index peaked in March of 2000.

cluster on the coasts. The east or west coast is home to 72% of information technology firms, but only 42% of non-information technology firms are located on the coasts.

The average correlation of firms within the entire information technology industry is only modestly higher than the average correlation of firms within other 2-digit GICS industries, but subsectors of the information technology industry such as the 4-digit GICS industry semiconductor manufacturing and equipment have very high within-industry correlations. The average correlation of firms within the information technology industry is 0.43, while the average of this statistic for all 2-digit GICS industries is 0.37. The average level of correlation within semiconductor manufacturing and equipment firms is 0.72, higher than any other 4-digit GICS industry. The average level of this statistic for all 4-digit GICS industries is 0.42. Since we are looking for the relationship between correlation and distance, the combination of coastal clustering and above average correlation might complicate our analysis.

Information technology firms appear to be similar to non-information technology firms in most respects. Only the number of analysts is statistically different for the two types of firms, with information technology firms having significantly more analyst coverage than non-information technology firms (19.9 versus 15.2). Coverage of information technology firms is higher than of non-information technology firms even after controlling for market value and other characteristics.

Table 2 reports the number of S&P 500 firms and the percentage of the total market capitalization by state and 4-digit GICS industry.[6] Although 38 different states are represented in the index, New York and California contain 124 of the S&P 500 firms and account for more than one-third of the index's total market capitalization. More than 50% of the S&P's total market value is accounted for from just five states (New York, California, Texas, New Jersey, and Illinois).

[6]Note that we classify Washington DC as a state.

More than one third of the market capitalization of the entire S&P 500 is in 4 industries; pharmaceuticals, capital goods, diversified financial, and banking. The smallest industry, real estate, accounts for only 0.6% of the market capitalization of the S&P 500. Twenty-four different industries are represented at this level (4-digit) of the index.

3. Empirical Results

A. Plots of Average Correlation With Distance

As a preliminary investigation we look first at simple plots of stock return correlation and distance between firms. Figure 4 shows the average correlation of the monthly returns of S&P 500 firms over our sample period by 20 intervals of distance between the headquarters of the firms. Firms within approximately 100 miles of each other have, on average, correlation coefficients of around 0.24. This falls quickly to about 0.20, and then falls gradually to about 0.14 at a distance of 2,000 miles. At distances greater than 2,000 miles correlations are higher, apparently because coastal firms are correlated with each other.

It is possible that this "coastal effect" is due to the facts that information technology firms are more concentrated on the coasts than non-information technology firms, and that information technology firms are more correlated with each other during our sample period than firms in other industries. At distances greater than 2,000 miles, a disproportionate number of pairs of firms consist of two information technology firms, whose correlation is usually higher than two typical firms from different industries. On the other hand, some investors might focus on firms located on the coasts, and coastal firms might move together because of the trading activity of their common owners.

Figure 5 is identical to Figure 4, except that information technology firms are excluded. The pattern of correlation by distance is similar except that the coastal effect is eliminated. Figure 6 shows correlation between pairs of firms that are in the same 4-digit GICS industry. Again, the coastal effect is eliminated, indicating that the apparent effect is due to industry clustering. Correlation declines with distance up to approximately 1,300 miles, and at greater distances there is no clear pattern.

The distance effect that we observe in Figures 4-6 could be the result of industries clustering together geographically. If automobile firms are concentrated around Detroit, for example, then the fact that these nearby firms are highly correlated could have nothing to do with distance, and might only be due to the firms belonging to a common industry. In the regression analysis that follows we attempt to control for industry effects.

B. Regression Analysis

The graphs in the previous section suggest that correlation of stock returns declines with distance out to approximately 2,000 miles. At distances greater than 2,000 miles, correlation might increase with distance, but perhaps only for information technology firms. The question remains as to whether this decline is due to industry effects or other local characteristics, such as common ownership of nearby firms. To distinguish between these two theories it is very important to carefully define and control for industrial grouping of firms.

GICS industrial classifications, described above in Section 1, allow us to define the industry of individual firms very specifically. At the 8-digit level there are 118 different industries in our sample of 456 firms, meaning that there are, on average, fewer than four firms per industry. In the regression analysis that follows, for each possible pair of firms we control for the correlation between the 8-digit industries of the two firms.

Our regression results are reported in Table 3. Each observation in the regression models represents a pair of firms in the S&P 500. The dependent variable is the correlation between the monthly returns of the two firms over the 5-year time period 2000-2004. In other words, each observation represents an element of the variance-covariance matrix of firm returns. Only observations on the upper (or lower) triangle of this matrix are used, since the upper and lower triangles are mirror images. Elements on the diagonal are not used since all of them have correlation equal to one and distance equal to zero by definition.

If all of the S&P 500 firms were in our sample we would have $(500^2 - 500)/2$ or 124,750 observations. Excluding firms with missing data or location outside of the U.S. leaves us with 103,740 observations. Independent variables in the regressions represent different relationships between the two firms, in other words, between firms i and j in the variance-covariance matrix. These variables include the distance between the two firms and the difference in the market value and other characteristics of the two firms. The correlation of the 8-digit industries that the two firms belong to is another independent variable, as is a dummy variable indicating whether the firms belong to the same industry.

Since each firm is used to calculate variables in many different observations, the observations are clearly not independent, and so the estimated standard errors are biased downward. In OLS regressions, which we ran but do not tabulate in this paper, the independent variables appear to be more statistically significant than they actually are. We address this bias by using a bootstrap method.

Our bootstrap method is to select random samples of the S&P 500 firms, with replacement. Sampling with replacement means that a particular firm might appear multiple times, or not at all in a particular sample. Correlations between each possible pair of the selected firms are calculated and these are the observations for the regression. Note that when a firm does not appear in a sample, its correlations with all other firms are removed.

This sampling procedure is repeated 1,000 times, and a new regression is performed on each of the 1,000 samples. This results in 1,000 estimates of the regression coefficients, and allows the mean and standard deviation of these estimates to be calculated. Statistical significance is determined by a t-test on the ratio of the estimated means of the coefficients to their estimated standard deviations. The reported R-squared is the average of the R-squared from these 1,000 regressions.

This sampling method has been used in similar situations in other fields of research. For example, Burrows, Moore, and James (2002) examined the growth of barnacle populations on the Scottish coast, and found that the correlation of growth rates varied by the distance between the populations. They ran 500 regressions, each with observations obtained by sampling the different populations with replacement. Other studies (also in biology) using bootstrap methods in similar situations include Swanson and Johnson (1999) and Paradis (2000). Bootstrap methods are discussed in more detail in Davison and Hinkley (1997).

The t-statistic normally provides information on the reliability of results, given the size of the sample. One interpretation of this statistic is that it shows the likelihood of obtaining the same results if a number of random samples of the same size were obtained. The bootstrap method simulates the process of taking many random samples, and then directly obtains the distribution of the parameter estimates from the estimates calculated from each sample. This distribution is assumed to be representative of the distribution that would be obtained by taking different samples from the entire population.

The coefficient estimates from the bootstrap regressions shown in Table 3 are very similar to those obtained using OLS regression, not tabulated in this paper, since the lack of independence of the observations biases only the standard errors, not the coefficient estimates. The reported t-statistics for many variables in Table 3 are considerably lower than the OLS t-statistics.

Model 1 in Table 3 demonstrates that the correlation between two firms is strongly influenced by the correlation of the two industries to which the firms belong. Two firms that are not in the same industry and whose respective industries are completely uncorrelated would be expected to have a correlation coefficient of 0.000, the intercept term reported in Model 1. If, on the other hand, two firms are in the same industry, their expected correlation coefficient would be calculated using the regression coefficients shown in equation 1. The correlation coefficient between firms i and j is given by C_{ij}, IC_{ij} is the correlation of the industry groups to which firms i and j belong, and D_{ij} is a dummy variable indicating whether firms i and j belong to the same industry.

$$C_{ij} = \alpha + \beta_1 IC_{ij} + \beta_2 D_{ij} \tag{1}$$

For two firms in the same industry, IC_{ij} would be equal to one because each industry is perfectly correlated with itself, and D_{ij} would also be equal to one, so the two firms would be expected to have a correlation coefficient of $\alpha + \beta_1 + \beta_2$. As reported in Table 3, Model 1, α is estimated to be equal to 0.000, β_1 is estimated to be equal to 0.596 and β_2 is estimated to be equal to be -0.131, so the expected correlation coefficient would be 0.465. The average correlation between 8-digit GICS industries is 0.31, so two average firms in different industries would be expected to have a correlation coefficient of $\alpha + \beta_1 0.31 + \beta_2 0.00$ or 0.185. We also include a variable for information technology firms, since these firms had an unusual amount of comovement during our sample period.

Model 2 adds information on the distance between firms. The model includes straight-line distance between firms to a distance of 2,000 miles, another variable for the distance between firms beyond 2,000 miles, and a dummy variable indicating whether the firms are further than 2,000 miles apart. These variables allow the slope and intercept of fitted lines to differ at distances greater and less than 2,000 miles.

The distance of 2,000 miles was chosen by taking the maximum of a series of Chow statistics as was suggested by Quandt (1960).[7] Distance up to 2,000 miles is statistically significant at the 3% level. The coastal effect is not statistically significant in Model 2, meaning that we cannot conclude that correlation increases with distance beyond 2,000 miles, but we can reject the hypothesis that the relationship between correlation and distance is the same closer to and further than 2,000 miles. The coefficient on the dummy variable indicating distance greater than 2,000 miles demonstrates that the intercept of the least squares line for firms greater than 2,000 miles apart is lower than for closer firms, and the coefficients on the distance variables indicates that the slope is higher. This is consistent with the evidence from the graphs in the previous section. The regressions of this section help to interpret the evidence from the graphs by indicating that the negative slope up to 2,000 miles is statistically significant, but the positive slope beyond 2,000 miles is not statistically significant. The two lines are, however, different from each other.

Model 3 adds various differences in characteristics between firms. We might expect that returns of firms that differ in various ways would be less likely to be highly correlated than firms that are similar. Differences in insider and institutional ownership, market value, and number of analysts are not statistically significant. The difference in the populations of headquarter cities is also statistically insignificant. This is in contrast to Pirinsky and Wang (2006), who find, surprisingly, that comovement of returns is higher in large cities.[8]

Beta differences have a large and a statistically significant effect on return correlation.

[7]See Hansen (2001) for more details. The maximum Chow statistic occurred at 1,980 miles.

[8]This is reported in Table VIII of Pirinsky and Wang (2006). The effect of city size is eliminated in Pirinsky and Wang (2006) when controls are added for MSA personal and investment income.

This is not a surprise, since firms that both are highly correlated with the overall market would be expected to be correlated with each other. The three betas that are used are those of Fama and French (1993); the stock market, the difference in the returns on portfolios of small and large firms, and the difference in the returns on portfolios of high and low book-to-market ratios. Including these variables in the model slightly increases the estimated effect of distance, now statistically significant at the 2% level.

To interpret the magnitude of these effects, consider two stocks 100 miles apart with standard deviations of monthly returns equal to σ_x and σ_y, each with a value of 0.11. The correlation coefficient, ρ_{xy}, between these two firms is 0.2. Now suppose that the firms are at the same location and that they have many common local owners who sometimes trade together. Each month there is a mean zero shock with standard deviation σ_s that causes these local owners to buy or sell, and this trading adds a common mean zero random term to both returns. Model 3 predicts that their correlation coefficient would be 0.0012 higher, or 0.2012. This new correlation coefficient of the returns of the two stocks, $\hat{\rho}_{xy}$ can be expressed as:

$$\hat{\rho}_{xy} = \frac{Cov(x,y) + \sigma_s^2}{\sigma_x \sigma_y + \sigma_s^2} \tag{2}$$

Solving for σ_s, we obtain:

$$\sigma_s^2 = \frac{\hat{\rho}_{xy}\sigma_x\sigma_y - Cov(x,y)}{1 - \hat{\rho}_{xy}} \tag{3}$$

Substituting in our assumed values of the original standard deviations and covariance of the stock returns and the new correlation coefficient, the standard deviation of the monthly shock to stock returns is approximately 43 basis points. In other words, an increase in the correlation coefficient of 0.0012 is equivalent to average monthly deviations of returns resulting from common investor activity of 43 basis points.

Forty three basis points per month is equivalent to 5.3% per year, which is higher than previous estimates of the advantage that investors earn from local stocks. For example, Coval and Moskowitz (2001) found an advantage of 2.67% per year and Ivkovic and Weisbenner (2005) report an advantage of 3.2%. If these movements are caused by local information or local trading activity, some local investors may be better able to predict these movements than non-local investors and earn extra returns. It is unlikely, however, that any local investors would be able to correctly anticipate all of these local movements, so it seems reasonable that our estimate of total local variance is higher than previous estimates of the advantage earned by local investors.

C. Robustness Tests

Our results do not appear to be driven by clusters of narrowly defined industries, but it is possible that we have missed the effect of clusters of related industries as described by Enright (2003) and Porter (2000). To test this possibility we included additional variables for the correlation of the 2, 4, and 6-digit industries that firms belong to. These results are reported in Table 4. The estimates and the standard errors of the estimates of the effect of distance are nearly unchanged. Correlation at the 8-digit level is by far the most important of the different industrial classifications, but correlation at the 2-digit industry provides some additional explanatory power.

Perhaps the most interesting change resulting from including additional industry information is that the positive relationship between correlation and distance when firms are separated by more than 2,000 miles is now statistically significant at approximately the 5% level. This result provides some weak evidence that, even after controlling for industrial grouping, coastal firms are correlated with each other in a similar manner as nearby firms. It is possible that there are bi-coastal investors who concentrate on firms located on the east and west coasts, and that their trading activity affects returns of these firms.

The most prominent event during our sample period of 2000-2004 was the terrorist attack on New York City and Washington D.C. on September 11, 2001. This event could have affected our estimate of the relationship between stock return correlation and distance, since investors might have worried about the effect of future attacks on certain large cities. If this were the case, then many close firms in large cities might have had simultaneous negative shocks to their returns, resulting in high measured correlation for nearby firms. We re-estimated our models eliminating return data for the 6 months beginning September, 2001. This change in the sample period had a trivial effect on the magnitude and statistical significance of the coefficients of the models.

We also attempted to test whether the relationship between correlation and distance is related to the market value of firms. Interaction terms obtained by multiplying the distance between firms and the difference in the size of the firms were not statistically significant in our regressions. We tried a number of specifications, including difference in logs, with the same results. It seems reasonable to expect that the effects of local ownership would be stronger for smaller firms, as reported in Pirinsky and Wang (2006), but we were unable to find any evidence of this in our sample.

It is possible that location in the same state increases the correlation of the returns of two firms. To test for this possibility, a dummy variable indicating whether a pair of firms were located in the same state was included in the regression analysis. The magnitude of the coefficient and statistical significance of distance was unaffected, and the dummy variable was not statistically significant.

D. Nonparametric Regression

In this section we relax our previous assumption that the relationship between correlation and distance is linear. The regression analysis of the previous section assumes a linear relationship between correlation and distance, but we have no theoretical basis for making this assumption, even in an idealized world. Geographic peculiarities such as the lack of firms in the western mountains/deserts are likely to make the relationship more complicated. The regression analysis of the previous section attempted to deal with nonlinearities with dummy variables for small and large distances and by allowing the slope to differ at distance greater than and less than 2,000 miles.

Another method of estimating relationships with unknown functional form is nonparametric regression. The basic idea of this method is to fit a curve to the data that balances two objectives: minimization of prediction error and minimization of curvature. Curvature could be minimized by fitting a straight line to the data, in other words, an ordinary least squares regression. Prediction error could be minimized by fitting a "bumpy" curve which wiggled wildly in order to intersect with each data point. By weighting the two objectives, a curve that is in between these two extremes can be generated. Hopefully, this method will allow us to filter out the noise from the data and gain some insight into the underlying relationship between return correlation and firm distance.

We begin with the residuals from Model 3 of Table 3. This model includes industry and firm difference information, but for this analysis we did not include the variables relating to distance. We modeled the residuals from this regression using a least squares cubic spline. If d_i is the distance between firms for observation i, and r_i is the residual for observation i, the cubic spline curve, $s(d_i)$ minimizes the following expression:

$$p\Sigma_i \left(r_i - s(x_i)\right)^2 + (1-p)\int \frac{d^2 s}{dx^2} dx \tag{4}$$

The smoothness of the curve is determined by the choice of the smoothing parameter p. If p is equal to 0 the curve will simply be an ordinary linear least squares fit. If p is equal to 1, the curve will be forced to go through each point, although it will still be a continuous, differentiable curve. With very noisy data and p equal to 1, the curve will wiggle through each data point. As p is reduced from 1, the curve smoothes out, ignores noise, and, hopefully, reveals the underlying relationship (Davison and Hinkley 1997).

To select the smoothing parameter, we used a technique known as "leave-one-out-cross-validation." In this technique, the curve is estimated once for each observation in the data set. Each time, one observation is left out and the prediction error for that point is calculated. The squared errors of each of these estimations is then summed. We repeated this process for the smoothing parameters 0.1 through 0.9 and found that a parameter of 0.6 minimized the sum of squared out-of-sample prediction errors. This parameter seems to minimize spurious fluctuation while preserving the overall structure of the data.

We took 10,000 random samples of 1,000 points of the data, computed a cubic spline for each sample, and then computed the mean and 90% confidence interval of the splines at each distance, using a smoothing parameter of 0.6. Figure 7 shows a plot of the mean and confidence interval by distance. The downward slope for distances up to 1,000 miles is apparent, but the confidence interval for longer distances widens considerably. The slope appears to be steeper for very short distances, but again, the confidence interval is wider at these distances.

These results are consistent both with the simple graphs and the regression analysis presented in Section 3. All three analyses show clear evidence that, up to a distance between 1,000 and 2,000 miles, correlation of monthly stock returns declines with distance. All three also show weaker evidence that correlation increases with distance beyond 2,000 miles.

4. Conclusion

The S&P 500 is a collection of large, well-known, widely-held firms. Tests of the effects of geographic location on return correlation would seem likely to fail using this sample. We find, however, that the monthly returns of S&P 500 firms with headquarters that are geographically close are more highly correlated than firms that are distant from each other.

Our primary result is that the correlation of two S&P firms is larger, the closer the two firms are geographically from each other. Our bootstrap regression results show that the correlation coefficient between two stocks increases by 12 basis points for every 100 mile reduction in distance, an increase which is consistent with local shocks to stock returns of 43 basis points per month. We also find weak evidence that firms on opposite coasts are more highly correlated than other firms. It is possible that some investors specialize in coastal firms and their trading activity influences the returns of these firms in a similar manner as local investors affect the returns of nearby firms. This effect may be particularly concentrated in information technology firms.

A previous investigation of the relationship between location and stock return correlation, Pirinsky and Wang (2006), suffered from methodological problems, and did not find evidence of an effect for large firms. Their results are biased by non-independent observations and sub-industry correlation, which we counter by using bootstrap regressions and finely detailed industry definitions. Our analysis finds a relationship between distance and return correlation for the largest publicly traded firms in the United States.

We believe that our findings are consistent with a world in which firms are disproportionately held by local investors who tend to trade together. Local events affect the buy/sell decisions of investors, and trading activity in all nearby firms is affected by these events, causing return correlation. Previous research has found evidence from a variety of sources that investors are biased in favor of owning local stocks, but an implication of this finding, that stock return correlation will be related to distance, had not, with the exception of Pirinsky and Wang (2006), been previously tested. Our results strengthen the case that local investor bias is a widespread and important phenomenon.

References

Audretsch, D., Feldman M. 1996a. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review*, 86, 630–640.

Audretsch, D., Stephan P. 1996b. Company-Scientist Locational Links: The Case of Biotechnology. *American Economic Review*, 86, 641–652.

Bhojraj, S., Lee C. Oler D. 2003. What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research*, 41, 745–774.

Brown, J., Ivkovic Z. Smith P. Weisbenner S. 2004. The Geography of Stock Market Participation: The Influence of Communities and Local Firms. Finance and Economics Discussion Series, Federal Reserve Board, paper 2004-22.

Burrows, M., Moore J. James B. 2002. Spatial Synchrony of Population Changes in Rocky Shore Communities in Shetland. *Marine Ecology Progress Series*, 240, 39–48.

Coval, J., Moskowitz T. 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. *Journal of Finance*, 54, 2045–2073.

Coval, J., Moskowitz T. 2001. The Geography of Investment: Informed Trading and Asset Prices. *Journal of Political Economy*, 109, 811–841.

Davison, A., Hinkley D. 1997. *Bootstrap Methods and their Applications*. Cambridge: Cambridge University Press, 1997.

Enright, M. 2003. Regional Clusters: What We Know and What We Should Know. *Innovation Clusters and Interregional Competition*, 99–129. Brocker, J., Dohse, D., Soltwedel, R. Eds.

Fama, E., French K. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3–56.

Fama, E., French K. 1997. Industry costs of equity. *Journal of Financial Economics*, 43, 153–193.

Glenn, E., Glaeser E. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105, 889–927.

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

Hansen, B. 2001. The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity. *Journal of Economic Perspectives*, 15, 117–128.

Hong, H., Kubrik J. Stein J. 2003. Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers. Working paper, Princeton University.

Huberman, G. 2001. Familiarity Breeds Investment. *Review of Financial Studies*, 14, 659–680.

Ivkovic, Z., Weisbenner S. 2005. Local does as Local is: Information Content of the Geography of Individual Investors Common Stock Investments. *Journal of Finance*, 60, 267–306.

Loughran, T., Schultz P. 2004. Weather, Stock Returns, and the Impact of Localized Trading Behavior. *Journal of Financial and Quantitative Analysis*, 39, 343–364.

Loughran, T., Schultz P. 2005. Liquidity: Urban versus Rural Firms. *Journal of Financial Analysis*, 78, 341–374.

Marshall, A. 1920. *Principles of Economics*. London: Macmillan (8th edition).

Paradis, E. 2000. Spatial Synchrony in Populations of Birds: Effects of Habitat, Population Trend, and Spatial Scale. *Ecology*, 81, 2112–2125.

Pirinsky, C., Wang Q. 2006. Does Corporate Headquarters Location Matter for Stock Returns? *Journal of Finance*, 61, 1991–2015.

Porter, M. 2000. Locations, Clusters, and Company Strategy. *The Oxford Handbook of Economic Geography*, 253–274. Clark, G., Feldman, M., Gertler, M. Eds.

Quandt, R. 1960. Tests of the Hypothesis that a Linear Regression Obeys Two Separate Regimes. *Journal of the American Statistical Association*, 55, 324–330.

Swanson, B., Johnson D. 1999. Distinguishing Causes of Intraspecific Synchrony in Population Dynamics. *Oikos*, 86, 265–274.

Zhu, N. 2002. The Local Bias of Individual Investors. *Working Paper, Yale University.*

Table 1: **Descriptive Statistics.** For each item except correlation, a mean is calculated for each firm over the time period 2000-2004. The statistics reported here are calculated across firms. All S&P 500 firms are included except those with missing data or that have operating headquarters outside of the United States. Information technology firms are defined as belonging to industry 45 of the Global Industry Classification Standard (GICS).

Panel A: All Firms

	Mean	Median	St. Dev.	Min	Max	N
Market Capitalization ($ bill.)	19.21	8.22	36.94	0.94	313.67	456
Insider Ownership (%)	4.84	1.12	8.29	0.03	70.83	456
Institutional Ownership (%)	74.30	76.25	14.57	5.80	96.20	456
Average Number of Analysts	16.00	15.40	6.39	1.24	38.03	456
Population of HQ City (mill.)	5.84	3.59	6.03	0.01	18.64	456
Monthly Return (%)	1.20	1.21	1.09	-2.97	5.68	456
Correlation of Returns (%)	19.47	19.71	19.05	-56.16	88.78	103,740
Distance (thousands of miles)	1.06	0.73	0.17	0.00	2.72	103,740

Panel B: Information Technology Firms

	Mean	Median	St. Dev.	Min	Max	N
Market Capitalization ($ bill.)	20.64	7.00	43.54	1.06	288.79	74
Insider Ownership (%)	4.48	1.29	6.69	0.03	28.83	74
Institutional Ownership (%)	73.78	76.25	15.08	29.10	96.20	74
Average Number of Analysts	19.88	20.16	6.92	6.87	38.03	74
Population of HQ City (mill.)	4.85	2.42	5.11	0.01	18.64	74
Monthly Return (%)	0.65	0.46	1.36	-2.97	4.58	74
Correlation of Returns (%)	42.61	43.77	17.71	-17.23	88.78	2,701

Panel C: Non-Information Technology Firms

	Mean	Median	St. Dev.	Min	Max	N
Market Capitalization ($ bill.)	18.87	8.27	35.53	0.94	313.67	382
Insider Ownership (%)	4.90	1.08	8.58	0.04	70.83	382
Institutional Ownership (%)	74.55	76.30	14.47	5.80	96.20	382
Average Number of Analysts	15.21	14.68	6.03	1.24	32.51	382
Population of HQ City (mill.)	6.04	4.16	6.18	0.01	18.64	382
Monthly Return (%)	1.30	1.28	0.97	-1.47	4.22	382
Correlation of Returns (%)	20.59	20.79	18.11	-46.56	88.49	72,771

Table 2: **State and Industry Distribution.** State is location of operational headquarters of firms. Industry definitions are from the Global Industry Classification Standard (GICS) at the 4-digit level. An asterix (*) denotes 4-digit industries that are within the 2-digit industry of information technology. All S&P 500 firms are included except those with missing data or that have operating headquarters outside of the United States.

State	Firms	% Mkt Cap	Industry	Firms	% Mkt Cap
California	69	14.38	Capital Goods	32	8.77
New York	55	20.29	Utilities	31	2.89
Texas	37	9.19	Materials	30	2.94
Illinois	32	5.16	Healthcare	30	3.83
Ohio	26	3.94	I.T. Hardware*	30	7.25
Pennsylvania	21	1.86	Energy	28	6.76
New Jersey	19	6.59	Retailing	28	4.23
Massachusetts	17	1.73	Banks	27	7.46
Georgia	13	4.25	Software*	25	6.39
Minnesota	13	3.42	Consumer Dur., Apparel	21	1.15
North Carolina	13	3.47	Diversified Financial	21	7.80
Connecticut	11	4.97	Pharmaceuticals	20	10.91
Michigan	11	1.46	Semiconductors*	19	3.87
Florida	10	0.55	Insurance	18	4.10
Tennessee	10	0.91	Food, Bev., Tobacco	17	4.51
Virginia	9	1.35	Media	13	2.64
Washington	9	4.45	Consumer Services	12	1.24
Maryland	8	0.63	Commercial Serv., Supp.	10	0.84
Wisconsin	8	0.80	Food, Staples Retailing	9	4.16
Colorado	7	0.78	Transportation	8	1.26
Missouri	7	1.08	Real Estate	8	0.59
Indiana	6	1.28	Telecomunications	7	3.38
Alabama	5	0.35	Automobiles & Parts	6	0.62
Arkansas	5	3.04	Household, Pers. Prod.	6	2.40
Arizona	4	0.24			
Kentucky	4	0.27			
Rhode Island	4	0.30			
Delaware	3	0.84			
Louisiana	3	0.23			
Oklahoma	3	0.25			
Iowa	2	0.05			
Idaho	2	0.22			
Nebraska	2	0.33			
Nevada	2	0.16			
Oregon	2	0.14			
Washington DC	2	0.94			
New Hampshire	1	0.03			
Utah	1	0.06			
Totals	456	100.00		456	100.00

Table 3: **Bootstrap Regression Results.** Observations are pairs of S&P 500 firms. Dependent variable is the correlation coefficient of the two firms. Numbers in parentheses are t-statistics.

	Model 1	Model 2	Model 3
Intercept	0.000	0.011	0.074
	(0.06)	(8.59)	(39.24)
Industry Correlation	0.596	0.591	0.562
	(37.45)	(38.31)	(35.71)
In Same Industry	-0.131	-0.130	-0.124
	(-6.86)	(-7.00)	(-6.55)
Information Tech.	-0.009	-0.008	0.018
	(-1.11)	(-0.93)	(1.57)
2,000 miles apart		-0.088	-0.094
		(-1.67)	(-1.89)
Distance $<$ 2,000		-0.012	-0.012
		(-2.27)	(-2.49)
Distance $>$ 2,000		0.033	0.037
		(1.53)	(1.76)
Mkt. Value Difference			0.003
			(1.38)
Analyst Diff.			-0.003
			(-1.70)
Inst. Ownershp Diff.			-0.041
			(-1.58)
Insider Own. Diff			-0.037
			(-1.22)
Population Diff			-0.001
			(-0.24)
Alpha Diff.			0.016
			(0.05)
Beta1 Diff.			-2.581
			(-3.86)
Beta2 Diff.			-4.055
			(-4.93)
Beta3 Diff.			-1.261
			(-1.96)
Adjusted R^2	0.460	0.460	0.482
Observations	103,740	103,740	103,740

Table 4: **Bootstrap Regression Results.** Observations are pairs of S&P 500 firms. Dependent variable is the correlation coefficient of the two firms. Numbers in parentheses are t-statistics.

	Model 1	Model 2	Model 3
Intercept	-0.041	-0.030	0.032
	(-25.39)	(-17.13)	(13.76)
2-Digit Ind. Corr.	0.072	0.072	0.080
	(2.71)	(2.70)	(3.01)
4-Digit Ind. Corr.	0.042	0.041	0.032
	(1.33)	(1.31)	(1.02)
6-Digit Ind. Corr.	-0.013	-0.012	-0.023
	(-0.42)	(-0.41)	(-0.77)
8-Digit Ind. Corr.	0.547	0.542	0.522
	(22.35)	(22.39)	(21.58)
In Same 2-Dig. Ind.	-0.025	-0.025	-0.024
	(-2.66)	(-2.54)	(-2.55)
In Same 4-Dig. Ind.	-0.008	-0.008	-0.003
	(-0.82)	(-0.78)	(-0.32)
In Same 6-Dig. Ind.	0.045	0.043	0.040
	(3.48)	(3.12)	(3.02)
In Same 8-Dig. Ind.	-0.156	-0.154	-0.150
	(-7.77)	(-7.67)	(-7.52)
Information Tech.	0.004	0.005	0.032
	(0.41)	(0.55)	(2.81)
2,000 miles apart		-0.096	-0.105
		(-1.84)	(-2.10)
Distance $<$ 2,000		-0.012	-0.012
		(-2.27)	(-2.54)
Distance $>$ 2,000		0.036	0.041
		(1.65)	(1.94)
Size Difference			0.004
			(1.60)
Analyst Diff.			-0.003
			(-1.70)
Inst. Ownershp Diff.			-0.038
			(-1.49)
Insider Own. Diff			-0.038
			(-1.29)
Population Diff			-0.000
			(-0.03)

(Continued on next page.)

Table 4: **Bootstrap Regression Results.** (Continued from previous page.)

	Model 1	Model 2	Model 3
Alpha Diff.			0.040
			(0.15)
Beta1 Diff.			-2.441
			(-3.86)
Beta2 Diff.			-4.066
			(-4.85)
Beta3 Diff.			-1.477
			(-2.29)
Adjusted R^2	0.468	0.468	0.489
Observations	103,740	103,740	103,740

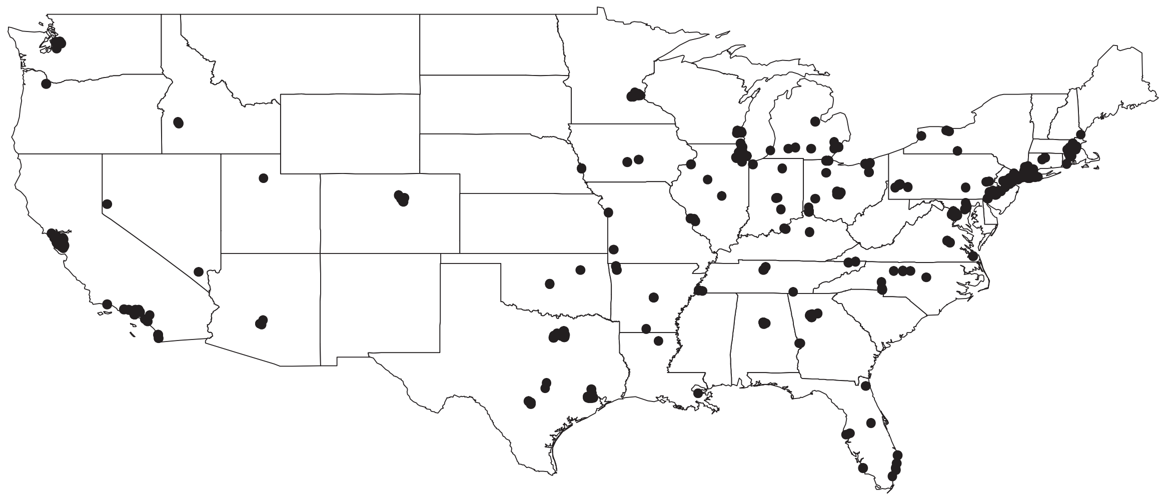

Figure 1: **Firm Location.** Dots represent the operational headquarters of firms in our sample. All S&P 500 firms are included except those with missing data or that have operating headquarters outside of the United States.

Figure 2: **IT Firm Location.** Location of information technology S&P 500 firms in our sample. Information technology firms are defined as belonging to industry 45 of the Global Industry Classification Standard (GICS).

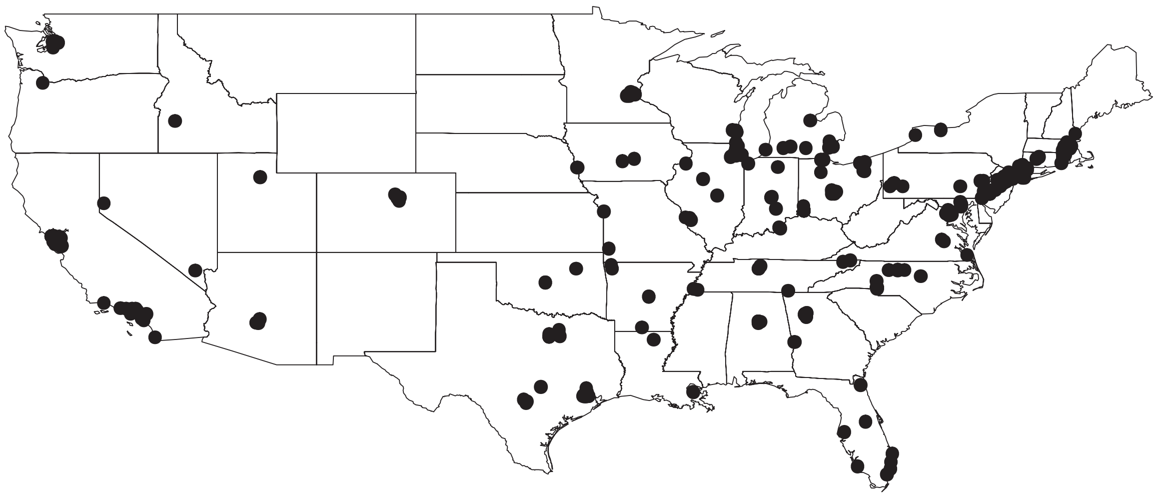

Figure 3: **Non-IT Firm Location.** Dots represent the operational headquarters of non-information technology S&P 500 firms in our sample. Information technology firms are defined as belonging to industry 45 of the Global Industry Classification Standard (GICS).

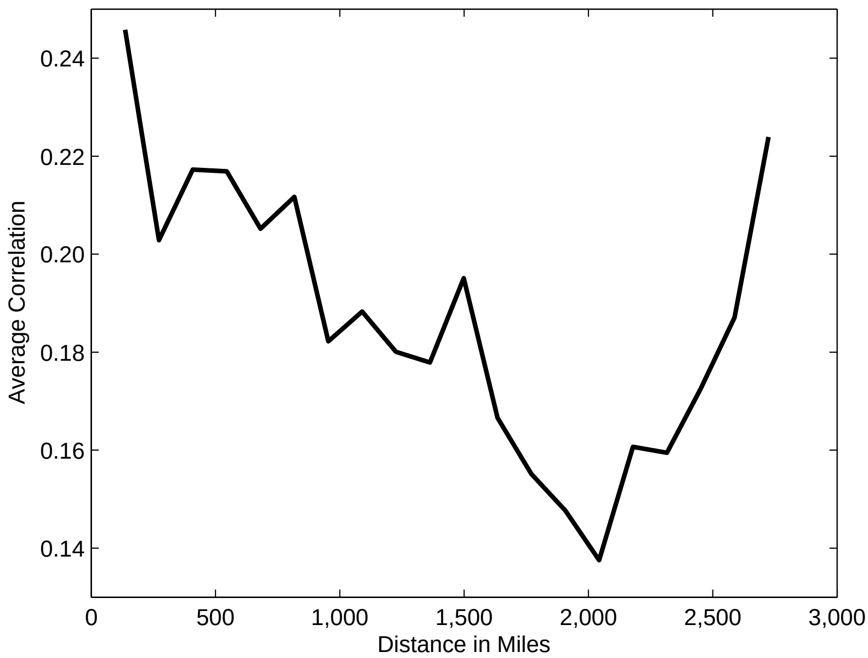

Figure 4: **Correlation vs. Distance.** Correlation coefficients of monthly stock returns of S&P 500 firms in our sample, 2000-2004, and distance between operational headquarters of the firms.

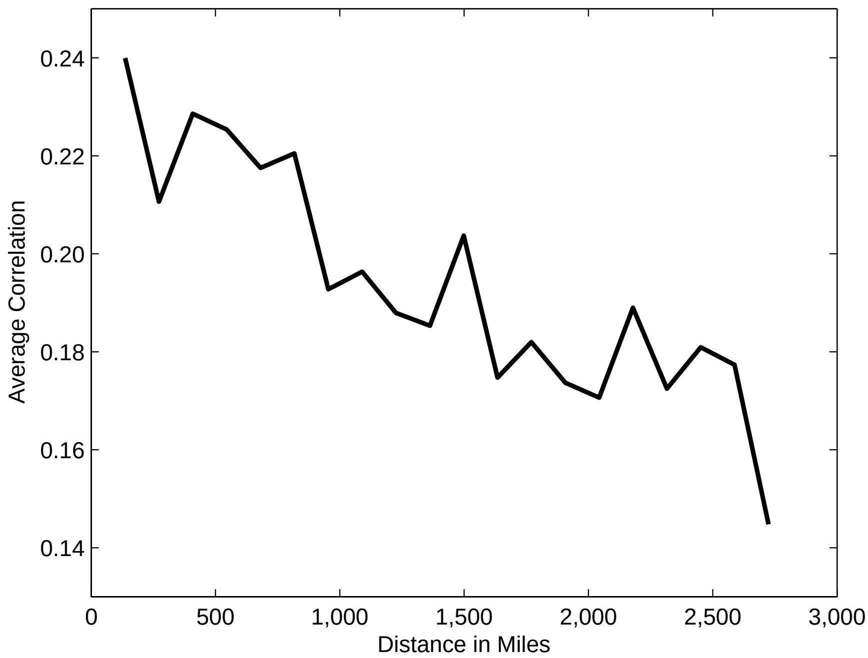

Figure 5: **Correlation vs. Distance for Non-IT Firms.** Correlation coefficients of monthly stock returns of non-information technology S&P 500 firms in our sample, 2000-2004, and distance between operational headquarters of the firms. Information technology firms are defined as belonging to industry 45 of the Global Industry Classification Standard (GICS).

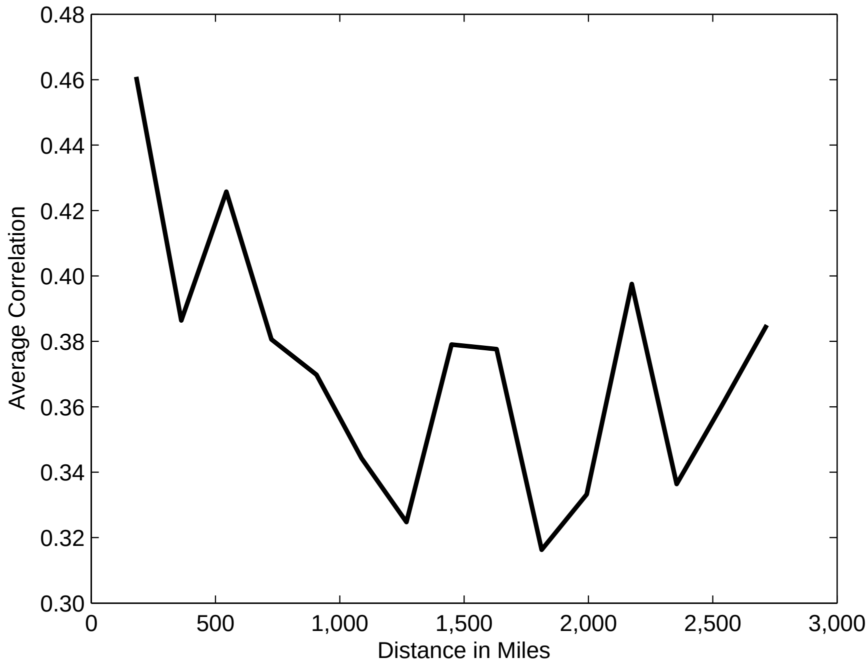

Figure 6: **Correlation vs. Distance for Firms in the Same Industry.** Correlation coefficients of monthly stock returns of S&P firms in our sample, 2000-2004, in the same 4-digit GICS industry and distance between operational headquarters of the firms.

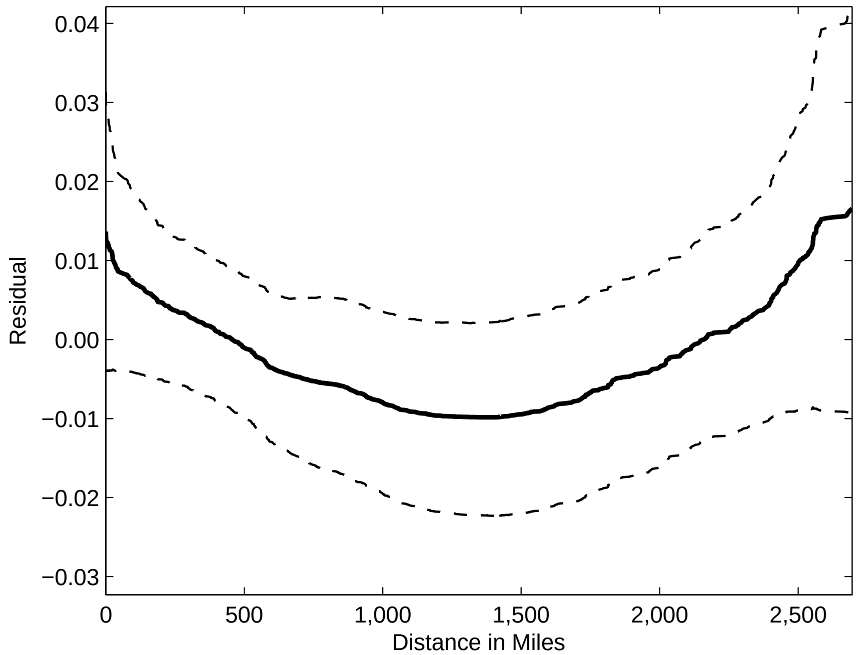

Figure 7: **Cubic Spline Fit of Residuals of Model 3, Table 3.** Solid line is the cubic spline fit of the residuals of Model 3 of Table 3 without distance variables. Dotted lines are 90% confidence intervals.