Comovement

Citation	Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler. 2005. "Comovement." Journal of Financial Economics 75 (2) (February): 283–317. doi:10.1016/j.jfineco.2004.04.003.
Published Version	doi:10.1016/j.jfineco.2004.04.003
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:27867240
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Comovement

Nicholas Barberis, Andrei Shleifer, and Jeffrey Wurgler
University of Chicago, Harvard University and New York University

March 2002*

Abstract

A number of studies have identified patterns of positive correlation of returns, or comovement, among different traded securities. We distinguish three views of such comovement. The traditional "fundamentals" view explains the comovement of securities through positive correlations in the rational determinants of their values, such as cash flows or discount rates. "Category-based" comovement occurs when investors classify different securities into the same asset class and shift resources in and out of this class in correlated ways. A related phenomenon of "habitat-based" comovement arises when a group of investors restricts its trading to a given set of securities, and moves in and out of that set in tandem.

We present models of each of the three types of comovement, and then assess them empirically using data on stock inclusions into and deletions from the S&P 500 index. Index changes are noteworthy because they change a stock's category and investor clientele (habitat), but do not change its fundamentals. We find that when a stock is added to the index, its beta and R-squared with respect to the index increase, while its beta with respect to stocks outside the index falls. The converse happens when a stock is deleted. These results are broadly supportive of the category and habitat views of comovement, but not of the fundamentals view. More generally, we argue that these non-traditional views may help explain other instances of comovement in the data.

*We thank Will Goetzmann, Mike Ryngaert, seminar participants at Columbia University, Harvard University, the London School of Economics, New York University, Rice University, UCLA, the University of Florida at Gainesville and Yale University for helpful comments, Huafeng Chen and Bill Zhang for helpful comments and outstanding research assistance, and Rick Mendenhall and Standard and Poor's for providing data.

1 Introduction

Researchers studying the structure of asset returns have uncovered numerous patterns of comovement. There is a strong common factor in the returns of small-cap stocks, for example, and also in the returns of value stocks, closed-end funds, stocks in the same industry, and bonds of the same rating and maturity. There is common movement within national markets and across international markets.

Common factors such as these have attracted considerable attention because of the possible role assets' loadings on them play in explaining *average* rates of return. However, little work has been done on understanding why the common factors arise in the first place. Why do certain groups of assets comove while others do not? What determines loadings, or betas, on these common factors? In this paper, we consider three theories of comovement – one traditional, two more novel – and present new evidence in support of the non-traditional theories.

The traditional view, derived from economies without frictions and with rational investors, is that comovement in prices reflects comovement in fundamental values. Since, in a frictionless economy with rational investors, prices equal fundamental value – in other words, the sum of an asset's rationally forecasted cash flows, discounted at a rate appropriate for their risk – any comovement in prices must be due to comovement in fundamentals.

An asset's fundamental value can change either because rational investors revise their expectations about future cash flows or because they apply a different discount rate to those cash flows. Under the traditional view, then, correlation in returns is due either to correlated changes in rationally expected cash flows or to correlated changes in rationally applied discount rates. Correlated discount rates can in turn arise because of news about interest rates or risk aversion, which affects all discount rates simultaneously, or because of correlated changes in assets' rationally perceived risk. There is little doubt that this "fundamentals" view of comovement explains many instances of common factors in returns: stocks in the oil industry move together because there is a common component to news about their future earnings, while the market factor in stock returns is at least in part due to changes in interest rates.[1]

A number of recent papers, however, present evidence suggesting that the traditional view of comovement is incomplete. Froot and Dabora (1999) study Siamese-twin stocks, which are claims to the same cash-flow stream, but are traded in different locations. Royal Dutch, traded primarily in the U.S., and Shell, traded primarily in the U.K., are perhaps

[1] The findings of Shiller (1989) illustrate the importance of accounting for changes in discount rates when examining patterns of comovement. He shows that the U.S. and U.K. stock markets comove more than can be explained by correlation in news about dividends alone; however, he also shows that allowing for plausible changes in discount rates can potentially explain the residual comovement.

the best known example. If return comovement is purely a reflection of comovement in news about fundamentals, these two stocks should be perfectly correlated. In fact, as Froot and Dabora show, Royal Dutch comoves more with the S&P 500 index of U.S. stocks than Shell does, while Shell comoves more with the FTSE index of U.K. stocks.

Hardouvelis, La Porta, and Wizman (1994) and Bodurtha, Kim, and Lee (1995) uncover related evidence in their studies of closed-end country funds, whose assets trade in a different location from the funds themselves. Since funds and their underlying assets represent claims to similar cash-flow streams, the fundamentals view of comovement predicts that fund returns and returns on their net asset values should be highly correlated. In fact, closed-end country funds comove much more with the national stock market in the country where they are traded than with the national stock market in the country where their *assets* are traded. For example, a closed-end fund invested in German equities but traded in the U.S. typically comoves more with the U.S. stock market than with the German stock market.

Fama and French (1995) investigate whether the strong common factors detected in the returns of value stocks and small stocks by Fama and French (1993) can be traced to common factors in the earnings of these stocks. While they do uncover a common factor in the earnings of small stocks, as well as in the earnings of value stocks, these cash-flow factors line up poorly with the return factors. Once again, there appears to be comovement in returns that has little to do with comovement in news about fundamentals.

Finally, Pindyck and Rotemberg (1990) find strong comovement in the prices of seven commodities – wheat, cotton, copper, gold, crude oil, lumber, and cocoa – that are chosen to be as independent of one another as possible. They are neither complements nor substitutes, are grown in different climates and are used for different purposes. Under the traditional view of comovement, the only plausible source of price correlation is news about aggregate demand. However, even after experimenting with a variety of forecasting models, Pindyck and Rotemberg are unable to find sufficient volatility in news about aggregate demand to fully explain the comovement.[2]

These examples suggest that investor trading patterns, and not just fundamentals, determine comovement. In this paper, we consider two specific models of such trading-induced comovement. The first model is based on the "category" view of comovement, recently analyzed by Barberis and Shleifer (2003). They argue that when making portfolio decisions, many investors first group assets into categories such as small-cap stocks, oil industry stocks, or junk bonds, and then allocate funds at the level of these various categories rather than at

[2]Pindyck and Rotemberg (1993) uncover similar evidence in an analogous study of stock returns. They construct groups of stocks that are in completely different lines of business and find that even though the stocks within each group are in different industries, their returns still comove strongly. This "excess" comovement remains after controlling for any cash-flow or discount rate correlation induced by news about future macroeconomic conditions.

the individual asset level. If some of the investors who use categories are noise traders with correlated sentiment, and if their trading affects prices, then as they move funds from one category to another, their coordinated demand will induce common factors in the returns of assets that happen to be classified into the same category, even if these assets' cash flows are largely uncorrelated.

Our second model of trading-induced comovement, which we refer to as the "habitat" view of comovement, starts from the observation that many investors choose to trade only a subset of all available securities. Such preferred habitats may arise because of transaction costs, international trading restrictions, or lack of information (Merton, 1987). As these investors' risk aversion or sentiment changes, they alter their exposure to the securities in their habitat, thereby inducing a common factor in the returns of these securities. For example, Lee, Shleifer, and Thaler (1991) argue that closed-end mutual funds are a preferred habitat of individual investors, and that therefore their market prices comove with the demand shifts of individual investors even when their fundamentals do not. More generally, this view of comovement predicts that there will be a common factor in the returns of securities that are held and traded by a specific subset of investors, such as individual investors.[3]

Trading-induced comovement is a simple way of understanding the empirical evidence described above. If small-cap stocks and value stocks form natural categories in investors' minds – and the large number of money managers and mutual funds focused on such stocks suggests that they do – then the category view of comovement predicts that there will be common factors in the returns of such stocks that are only weakly related to any common factors in their cash flows. Moreover, if many individual investors in the U.S. confine themselves to holding domestically traded securities, then the habitat view of comovement predicts that closed-end country funds traded in the U.S. will comove substantially with U.S. stocks even if their holdings consist of foreign equities.

The idea that trading unrelated to news about fundamental value might generate comovement builds on earlier evidence that such trading affects prices. Some of the best-known evidence of this type comes from stock index redefinitions. When an index is redefined, investors who follow it must reduce their holdings of securities that have been downweighted in the index and buy those whose weighting has increased. Under the efficient markets view, these demand shifts should not affect prices, as they carry no information about fundamental value. However, Harris and Gurel (1986), Shleifer (1986), and Lynch and Mendenhall (1997) find strong price effects for S&P 500 inclusions, while Kaul, Mehrotra, and Morck (1999) and Greenwood (2001) find similar effects in the Toronto Stock Exchange TSE 300 and Nikkei

[3]Other models which consider investor habitats are motivated by similar information and transaction cost considerations as our own, but focus on different issues. Merton (1987) analyses the cross-sectional implications when investors apply standard mean-variance analysis, but only over a subset of available assets. Our focus is on the effects of habitat-level demand shifts that affect all stocks in the habitat equally.

225 indices, respectively.[4]

In this paper, we return to the S&P 500 inclusion and deletion data. The same data that has proved useful in showing that uninformed demand can affect prices may also be helpful in showing that such demand can generate comovement. Since addition to the S&P 500 does not affect fundamental value, a stock's inclusion should not cause a change in the correlation of its fundamental value with the fundamental values of other stocks already in the index. Under the fundamentals view of comovement, then, it should not cause a change in the correlation of the stock's return with the return of the S&P. In particular, a univariate regression of a stock's return on the S&P return both before and after the stock's inclusion should produce similar slope coefficients, or S&P betas, and similar R^2s.

On the other hand, the vast popularity of S&P-linked investment products suggests that the index is a preferred habitat for some investors, and is viewed as a natural category by many more. Category-based investors include investors pursuing passive portfolio strategies through index funds as well as index arbitrageurs exploiting discrepancies between cash and futures prices. The trading-based theories therefore differ from the fundamentals view in their predictions about patterns of comovement before and after a stock's inclusion. In particular, simple models of the category and habitat views predict that in the univariate regression described above, the S&P beta and R^2 should increase after inclusion; that in a *bivariate* regression of a stock's return on both the S&P and a non-S&P "rest of the market" index, the S&P beta should rise after the stock's inclusion while the non-S&P beta should fall; that these patterns should go in the opposite direction for deletions; that these effects should be stronger in more recent data as the S&P becomes more widely used as a category and habitat; and that there should be a decrease in the correlation between S&P and non-S&P returns over time, again as the S&P becomes a more important category.

Our evidence supports the trading-based theories. Over a range of data frequencies, stocks added to the S&P increase their beta and R^2 with the S&P, while in bivariate regressions that control for non-S&P returns, increases in S&P beta are even more pronounced. Significant results in the opposite direction are observed when stocks are deleted from the index, and effects for both inclusions and deletions are stronger in more recent data. We also confirm a significant decrease in the correlation of S&P and non-S&P returns over time.

While adding a stock to the S&P 500 should not *cause* a change in the cash-flow covariance matrix, it is possible to construct alternative explanations for our results under which a stock's inclusion *coincides* with a shift in the covariance matrix. To rule these explanations out, we also conduct a "matching" analysis: for each "event" stock included into the S&P index, we search for a matching stock, drawn from the same industry as the event stock

[4]Numerous other papers present evidence consistent with uninformed demand affecting prices. These include French and Roll (1986), Lamont and Thaler (2000), Goetzmann and Massa (2001), Gompers and Metrick (2001), and Mitchell, Pulvino, and Stafford (2002).

and with similar market capitalization and recent growth in capitalization, but which is not added to the S&P. We find that at daily and weekly frequencies, the matching stocks display much smaller shifts in S&P and non-S&P betas than do the event stocks.

Our univariate regression results fit well with the evidence of Vijh (1994), who investigates whether the rise of S&P-linked products affects stocks' beta with respect to the *overall* market. He finds a significant increase in stocks' betas after inclusion, which is consistent with the increase in S&P beta we detect, given the dominant contribution of S&P stocks to the value-weighted market return.

In a recent paper, Greenwood and Sosner (2002) also test our model. Instead of focusing on the S&P 500, they use data on additions to and deletions from the Nikkei index. They find evidence of increases in beta and R^2 following a stock's addition to the index, and of decreases following deletions. Their evidence is thus also consistent with the predictions of our model; if anything, the results for the Japanese data are even stronger than those for the U.S. data.

In Section 2, we present some simple models illustrating the various views of comovement, as well as their distinct predictions. In Section 3, we test a number of these predictions using data on S&P 500 inclusions and deletions. Section 4 concludes.

2 Three Models of Comovement

In this section, we lay out three theories of return comovement. The models we present are simple, but they nevertheless allow us to illustrate the predictions of each theory. These predictions motivate the empirical work in Section 3.

In all three models, the economy contains a riskless asset in perfectly elastic supply and with a zero rate of return, and also $2n$ risky assets in fixed supply. Risky asset i is a claim to a single liquidating dividend $D_{i,T}$ to be paid at some later time T. The eventual dividend equals

$$D_{i,T} = D_{i,0} + \varepsilon_{i,1} + \ldots + \varepsilon_{i,T}, \tag{1}$$

where $D_{i,0}$ is known at time 0 and $\varepsilon_{i,t}$ becomes known at time t, and where

$$\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{2n,t})' \sim N(0, \Sigma_D), \text{ i.i.d over time.}$$

The price of a share of risky asset i at time t is $P_{i,t}$. The asset's return between time $t-1$ and time t is[5]

$$\Delta P_{i,t} \equiv P_{i,t} - P_{i,t-1}. \tag{2}$$

[5]For simplicity, we refer to the asset's change in price as its return.

2.1 Fundamentals-based Comovement

Under the fundamentals view, comovement in returns is due to comovement in news about fundamental value. This prediction emerges from a wide range of models. We present a simple example below. This model provides a natural benchmark that we can compare to our alternative models of comovement.

The economy contains a large number of identical agents known as "fundamental traders." They have CARA utility defined over the value of their invested wealth one period later, and take price changes to be normally distributed.[6] They therefore solve

$$\max_{N_t} E_t^F(-\exp[-\gamma(W_t + N_t{}'(P_{t+1} - P_t))]), \tag{3}$$

where

$$\begin{aligned} P_t &= (P_{1,t}, \ldots, P_{2n,t})' \\ N_t &= (N_{1,t}, \ldots, N_{2n,t})', \end{aligned}$$

and where $N_{i,t}$ is the number of shares allocated to risky asset i, γ governs the degree of risk aversion, E_t^F denotes fundamental trader expectations at time t, and W_t is time t wealth.

Optimal holdings N_t^F are given by

$$N_t^F = \frac{(V_t^F)^{-1}}{\gamma}(E_t^F(P_{t+1}) - P_t), \tag{4}$$

where

$$V_t^F \equiv \mathrm{var}_t^F(P_{t+1} - P_t),$$

with the F superscript in var_t^F again denoting a forecast made by fundamental traders.

If the total supply of the $2n$ assets is given by the vector Q, then given fundamental trader expectations about future prices, current prices satisfy

$$P_t = E_t^F(P_{t+1}) - \gamma V_t^F Q. \tag{5}$$

Rolling this equation forward and setting

$$E_{T-1}^F(P_T) = E_{T-1}^F(D_T) = D_{T-1},$$

where

$$D_t = (D_{1,t}, \ldots, D_{2n,t})',$$

leads to

$$P_t = D_t - \gamma V_t^F Q - E_t^F \sum_{k=1}^{T-t-1} \gamma V_{t+k}^F Q. \tag{6}$$

[6] This assumption is confirmed in equilibrium.

If fundamental traders set

$$V_t^F = \Sigma_D, \forall t, \tag{7}$$

equation (6) reduces to

$$P_t = D_t - (T - t)\gamma\Sigma_D Q. \tag{8}$$

This means that up to a constant

$$\Delta P_{t+1} = \Delta D_{t+1} = \varepsilon_{t+1}, \tag{9}$$

confirming fundamental traders' conjecture about the conditional covariance matrix of returns.

Equation (9) shows that in this economy, return comovement simply reflects comovement in news about fundamental value. More specifically, since discount rates are constant, it reflects comovement in news about future cash flows.[7] This model is useful for understanding many instances of common factors in returns. The strong market and industry factors in returns, for example, are at least in part due to market-level and industry-level factors in cash-flow news.

2.2 Category-based Comovement

Barberis and Shleifer (2003) argue that when making their portfolio decisions, many investors first group assets into categories based on some characteristic, and then allocate funds at the level of these categories rather than at the level of individual securities. Thinking about investments in terms of categories is particularly attractive to institutional investors who, as fiduciaries, must follow systematic rules in their portfolio allocation. Investing by category simplifies the investment process, and also provides a consistent way of evaluating the performance of money managers.

To test any predictions that emerge from a category-based model, it is important to have a concrete way of identifying categories. One place to start is to look at the labels mutual and pension fund managers use to describe their products to clients. If money managers are responsive to client needs, they will choose labels that correspond to the categories people like to use when thinking about investments. For example, since many money managers offer funds that invest in value stocks, "value stocks" may be a category in the minds of many investors. This way of thinking suggests that Treasury bonds, junk bonds, large stocks, small stocks, growth stocks, or stocks within a particular industry, country, or index are also all examples of categories.

[7] Discount rates are constant because the riskless rate is constant, as are investors' risk aversion and their perception of risk.

The category view of comovement holds that some of the investors who use categories are noise traders with correlated sentiment. As their sentiment changes, they channel funds in and out of the various categories. If these fund flows affect prices, they will generate common factors in the returns of assets that happen to be classified into the same category, even if these assets' fundamental values are uncorrelated. For example, if "value stocks" is a popular category, then as noise traders move funds in and out of value stocks in line with their changing sentiment about value stocks, they will create a common factor in value stock returns even if value stock earnings are completely uncorrelated.

To see this in a formal model, suppose that there are just two such categories, X and Y, and that risky assets 1 through n are in category X while assets $n+1$ through $2n$ are in Y. It may be helpful to think of X and Y as "old economy" and "new economy" stocks, respectively. We write noise trader demand $N_{i,t}^C$ for shares of an asset i in category X at time t as[8]

$$N_{i,t}^C = \frac{1}{n}\left[A_X + u_{X,t}\right],\ i \epsilon X \tag{10}$$

and for an asset j in category Y as

$$N_{j,t}^C = \frac{1}{n}\left[A_Y + u_{Y,t}\right],\ j \epsilon Y. \tag{11}$$

Here A_X and A_Y are constants, and $u_{X,t}$ and $u_{Y,t}$ are the time t shocks to noise trader sentiment about categories X and Y, respectively. They are distributed

$$\begin{pmatrix} u_{X,t} \\ u_{Y,t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_u^2 \begin{pmatrix} 1 & \rho_u \\ \rho_u & 1 \end{pmatrix}\right),\ \text{i.i.d. over time.}$$

The fact that the demand for all assets within a category is the same underscores the fact that these investors allocate funds at the category level and do not distinguish among assets in the same category.

This economy also contains fundamental traders whose objective function is the one in (3). In this case, they double up as market makers, treating the noise trader demand as a supply shock. Given their expectations about future prices, current prices are given by

$$P_t = E_t^F(P_{t+1}) - \gamma V_t^F(Q - N_t^C), \tag{12}$$

where

$$N_t^C = (N_{1,t}^C, \ldots, N_{2n,t}^C)'.$$

Rolling this equation forward, and setting $E_{T-1}^F(P_T) = D_{T-1}$, leads to

$$P_t = D_t - \gamma V_t^F(Q - N_t^C) - E_t^F \sum_{k=1}^{T-t-1} \gamma V_{t+k}^F(Q - N_{t+k}^C). \tag{13}$$

[8]The "C" superscript stands for Category.

We simplify this further by imposing a more specific structure on the cash-flow covariance matrix Σ_D. In particular, we suppose that the cash-flow shock to an asset has three components: a market-wide cash-flow factor which affects assets in both categories, a category-specific cash-flow factor which affects assets in one category but not the other, and a completely idiosyncratic cash-flow shock specific to a single asset. Formally, for $i \epsilon X$,

$$\varepsilon_{i,t} = \psi_M f_{M,t} + \psi_S f_{X,t} + \sqrt{(1 - \psi_M^2 - \psi_S^2)} f_{i,t}, \tag{14}$$

and for $j \epsilon Y$,

$$\varepsilon_{j,t} = \psi_M f_{M,t} + \psi_S f_{Y,t} + \sqrt{(1 - \psi_M^2 - \psi_S^2)} f_{j,t}, \tag{15}$$

where $f_{M,t}$ is the market-wide factor, $f_{X,t}$ and $f_{Y,t}$ are the category-specific factors, and $f_{i,t}$ and $f_{j,t}$ are idiosyncratic factors; ψ_M and ψ_S are constants which control the relative importance of the three components. Each factor has unit variance and is orthogonal to the other factors. This implies

$$\Sigma_D^{ij} \equiv \text{cov}(\varepsilon_{i,t}, \varepsilon_{j,t}) = \begin{cases} 1, \; i = j \\ \psi_M^2 + \psi_S^2, \; i, j \text{ in the same category, } i \neq j \\ \psi_M^2, \; i, j \text{ in different categories.} \end{cases} \tag{16}$$

In words, all assets have a cash-flow news variance of one, the pairwise cash-flow correlation between any two distinct assets in the same category is the same, and the pairwise cash-flow correlation between any two assets in different categories is also the same.

Now suppose that fundamental traders conjecture that the conditional covariance matrix of returns has the form

$$V_t^F = V = \sigma^2 \begin{pmatrix} A & B \\ B & A \end{pmatrix}, \; \forall t, \tag{17}$$

where

$$A = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_1 \\ \rho_1 & \cdots & \rho_1 & 1 \end{pmatrix}, B = \begin{pmatrix} \rho_2 & \cdots & \cdots & \rho_2 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho_2 & \cdots & \cdots & \rho_2 \end{pmatrix},$$

for some σ^2, ρ_1, and ρ_2.

Given this conjecture,

$$P_t = D_t - \gamma V(Q - N_t^C) - (T - t - 1)\gamma V(Q - A), \tag{18}$$

where

$$A = (\frac{A_X}{n}, \ldots, \frac{A_X}{n}, \frac{A_Y}{n}, \ldots, \frac{A_Y}{n})',$$

which means that up to a constant,

$$\Delta P_{t+1} = \varepsilon_{t+1} + \gamma V \Delta N_t^C. \tag{19}$$

This reduces to

$$\begin{aligned} \Delta P_{i,t+1} &= \varepsilon_{i,t+1} + \frac{\Delta u_{X,t+1}}{\phi_1} + \frac{\Delta u_{Y,t+1}}{\phi_2}, \; i\epsilon X, \\ \Delta P_{j,t+1} &= \varepsilon_{j,t+1} + \frac{\Delta u_{X,t+1}}{\phi_2} + \frac{\Delta u_{Y,t+1}}{\phi_1}, \; j\epsilon Y, \end{aligned} \tag{20}$$

where

$$\begin{aligned} \phi_1 &= \frac{1}{\gamma\sigma^2(\rho_1 + (1-\rho_1)/n)}, \\ \phi_2 &= \frac{1}{\gamma\sigma^2\rho_2}, \end{aligned} \tag{21}$$

confirming fundamental traders' conjecture about the structure of the conditional covariance matrix of returns: $\text{cov}(\Delta P_{i,t+1}, \Delta P_{j,t+1})$ is constant for all distinct assets i and j in the same category, and it is also constant for all assets i and j in different categories. We study equilibria in which the specific *values* of σ^2, ρ_1, and ρ_2 conjectured by fundamental traders are also confirmed by (20).[9]

Equation (20) shows that in this economy, there can be a common factor in the returns of a group of stocks simply because those stocks happen to belong to the same category. When noise traders experience a positive sentiment shock $\Delta u_{X,t+1}$ about category X, they invest more in *all* securities in X, pushing the prices of these assets up together.

The intuition for why $\Delta u_{X,t+1}$ affects the return on stock 1 is clear enough: when noise traders become bullish about old economy stocks, they channel funds into X, pushing the prices of all securities in that category up. Why $\Delta u_{Y,t+1}$ also affects the return on stock 1 is less obvious. Suppose that noise traders become bullish about new economy stocks, pushing up the prices of securities in Y. Fundamental traders, seeing an overvaluation, will short stocks in Y, and hedge themselves as much as possible against adverse fundamental news by buying stocks in X. In this way, the sentiment shock about category Y, $\Delta u_{Y,t+1}$, is also transmitted to stocks in X.

The fact that in our model, noise traders affect prices – and hence also, patterns of comovement – relies on the assumption that fundamental traders have horizons which end before cash-flow uncertainty is resolved at time T. If fundamental traders only cared about wealth at time T, they would be more aggressive in countering the effect of noise traders. However, since they have a one-period horizon, they are forced to worry about future noise trader demand, which makes them invest less aggressively. Equations (20) and (21) show that a high risk aversion γ or perceived stock volatility σ^2 make them particularly reluctant to bet against the noise traders, increasing the impact of the sentiment shocks on returns.

[9] It is straightforward to show that such equilibria exist for a wide range of values of the exogeneous parameters γ, ψ_M, ψ_S, σ_u^2, and ρ_u.

The idea that fundamental traders may have short horizons has been emphasized by earlier work on limits to arbitrage (De Long, Shleifer, Summers, Waldmann 1990, Shleifer and Vishny 1997). That such constraints might limit arbitrage capacity is supported by the considerable empirical evidence, cited in the introduction, suggesting that demand unrelated to news about fundamental value affects security prices. Moreover, Wurgler and Zhuravskaya (2001) confirm that arbitrageurs are particularly wary of countering noise traders when the risk of doing so is greater. They show that the price jump on inclusion into an index is much larger for stocks with poor substitutes, in other words, for those cases where arbitrageurs face higher risk.

To uncover evidence of category-induced comovement, we look for testable predictions that are unique to this economy. One set of predictions describes what happens when a stock enters a new category. Such reclassification can occur in many ways. For example, if the market capitalization of a large-cap stock declines sufficiently, it will enter the small-cap stock category. More simply, stocks are regularly added to indices like the S&P 500 and Russell 2000 to replace stocks that have been removed due to bankruptcy or merger.

Our first prediction is:

Proposition 1: Suppose that risky asset j, previously a member of Y, is reclassified into X. Then, assuming a fixed cash-flow covariance matrix Σ_D, and as the number of risky assets $n \to \infty$, the OLS esimate of β_j in the univariate regression

$$\Delta P_{j,t} = \alpha_j + \beta_j \Delta P_{X,t} + v_{j,t}, \tag{22}$$

where

$$\Delta P_{X,t} = \frac{1}{n} \sum_{l \epsilon X} \Delta P_{l,t}, \tag{23}$$

as well as the R^2 of this regression, increase after reclassification.[10]

The intuition is straightforward: when asset j enters category X, it is buffeted by noise traders' flows of funds in and out of that category. This increases its covariance with the return on category X, $\Delta P_{X,t}$, and hence also its beta loading on that return. For simplicity, we assume that the cash-flow covariance matrix remains fixed. A more general version of the proposition would predict that beta increases *more* than can be explained by any increase in cash-flow correlation.

A similar intuition lies behind the following prediction:

Proposition 2: Suppose that risky asset j, previously a member of Y, is reclassified into X. Then assuming a fixed cash-flow covariance matrix Σ_D, and as the number of risky assets

[10]Proofs of all propositions are in the Appendix.

$n \to \infty$, the OLS estimate of $\beta_{j,X}$ in the bivariate regression

$$\Delta P_{j,t} = \alpha_j + \beta_{j,X} \Delta P_{X,t} + \beta_{j,Y} \Delta P_{Y,t} + v_{j,t} \tag{24}$$

rises after reclassification, while the OLS estimate of $\beta_{j,Y}$ falls.

Proposition 2 identifies a test that is potentially more powerful than the test in Proposition 1. The essential prediction of the category view of comovement is that when a stock enters category X, it becomes more sensitive to the category X sentiment shock $\Delta u_{X,t}$. Of course, $\Delta P_{X,t}$ is not a clean measure of this sentiment shock; a substantial part of its variation comes from news about market-level cash flows, $f_{M,t}$. In regression (24), $\Delta P_{Y,t}$ can be thought of as a control for such news, making the coefficient on $\Delta P_{X,t}$ a cleaner measure of sensitivity to $\Delta u_{X,t}$.

Note that Propositions 1 and 2 will not hold if, as in Section 2.1., there are no noise traders with demand function (10) in the economy, or if fundamental traders are able to counteract their effect. In these cases, return correlation is completely determined by correlation in news about fundamental value. Therefore if, as assumed in the propositions, the cash-flow covariance matrix Σ_D remains constant, the correlation structure of returns will also remain constant. In other words, β_j and R^2 in Proposition 1 and $\beta_{j,X}$ and $\beta_{j,Y}$ in Proposition 2 will remain unchanged after reclassification.

One final prediction of the category view of comovement is:

Proposition 3: In the presence of noise traders with demand function (10), and as the number of risky assets $n \to \infty$, the correlation of the return on X with the return on Y,

$$\text{corr}(\Delta P_{X,t}, \Delta P_{Y,t}),$$

is lower than it would be in an economy that contains only fundamental traders.

When the economy contains only fundamental traders, the correlation of the returns of categories X and Y is completely determined by the correlation of the fundamentals of those two categories. Introducing noise traders adds less than perfectly correlated shocks to the returns of categories X and Y, lowering the correlation between them.

Proposition 3 becomes testable in the time series if the fraction of investors with demand functions in (10) grows over time; in that case, assuming a fixed cash-flow covariance matrix, the correlation of the two categories' returns should fall over time.

2.3 Habitat-based Comovement

The habitat view of comovement starts from the observation that many investors trade only a subset of all available securities. Such preferred habitats may arise because of transaction costs, international trading restrictions, or lack of information (Merton, 1987). For example, suppose that one group of investors – "habitat X" investors – trades only securities 1 through n, a set we again refer to as X, while another group – habitat Y investors – trades only $n+1$ through $2n$, set Y. We can think of assets 1 through n as U.S. stocks, and assets $n+1$ through $2n$ as U.K. stocks; there are many investors in both countries who restrict themselves to trading only domestic securities. We emphasize that X and Y play different roles here than in Section 2.2. There, they represent groups of assets that some investors do not distinguish between when allocating their demand. Here, they represent groups of assets that are the sole holdings of some investors.

Now suppose that habitat X investors experience an increase in risk aversion. They will then reduce their positions in all the risky assets they hold, generating a common factor in the returns of securities in X, even if those risky assets' fundamental values are uncorrelated. More generally, the habitat view of comovement predicts a common factor in the returns of any group of stocks that happens to be the primary holdings of a particular subset of investors.

To compare this view to the category-based view, suppose that habitat X investors' demand for risky assets is given by

$$\begin{aligned} N_{i,t}^{HX} &= \frac{1}{n}\left[A_X + u_{X,t}\right],\ i\epsilon X \\ N_{j,t}^{HX} &= 0,\ j\epsilon Y. \end{aligned} \tag{25}$$

We think of $u_{X,t}$ as tracking their level of risk aversion, changes in which lead them to alter their exposure to all assets in X. Of course, $u_{X,t}$ can also be interpreted as an indicator of sentiment about the future returns of assets in X, although the model does not require such an interpretation. By definition, habitat X investors' demand for assets in Y is zero.

Similarly, habitat Y investors' demand is

$$\begin{aligned} N_{i,t}^{HY} &= 0,\ i\epsilon X \\ N_{j,t}^{HY} &= \frac{1}{n}\left[A_Y + u_{Y,t}\right],\ j\epsilon Y. \end{aligned} \tag{26}$$

We assume

$$\begin{pmatrix} u_{X,t} \\ u_{Y,t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_u^2 \begin{pmatrix} 1 & \rho_u \\ \rho_u & 1 \end{pmatrix}\right),\ \text{i.i.d. over time.}$$

As before, we close the economy with fundamental traders who behave as in (3). Given their expectations about future prices, current prices are given by

$$\begin{aligned} P_t &= E_t^F(P_{t+1}) - \gamma V_t^F(Q - (N_t^{H1} + N_t^{H2})) \\ &= E_t^F(P_{t+1}) - \gamma V_t^F(Q - N_t^C), \end{aligned} \tag{27}$$

exactly as in (12). In other words, even though investors' demand functions are motivated differently here than in the case of category-based comovement, prices are the same. Once again, there will be a common factor in the returns of assets in X even if there is no common factor in their fundamentals.

The equivalence in equation (27) means that Propositions 1 through 3 also hold in this economy, with X and Y signifying investor habitats, not categories. For example, Proposition 1 should now be interpreted as predicting that if a stock becomes part of the habitat of a specific group of investors, it will comove more with the other assets in that habitat than it did before.

It is important to note that the habitat-based view of comovement depends on limits to arbitrage, just as the category-based view does. The fact that some investors trade only certain securities means that habitats X and Y can trade at different prices, even if their final cash flows are similar, thus opening up potentially attractive opportunities for unconstrained arbitrageurs. Since fundamental traders have short horizons in our model, they are unable to exploit these opportunities very aggressively.

3 Empirical Tests

Propositions 1 through 3 lay out predictions that hold in an economy where return comovement is in part due to category-based or habitat-based trading flows, but which do *not* hold in an economy where return comovement is entirely a function of comovement in news about fundamentals. We now test these predictions to see if we can uncover any evidence of trading-induced comovement.

To test the propositions, we need to identify a group of securities with three characteristics. First, the group must be viewed as a natural category, or must be a preferred habitat for many investors, or both. Second, since our first two propositions concern reclassification, there must be clear and identifiable changes in group membership over time. Finally, in order to control for fundamentals-based comovement, a security's inclusion or removal from the group should not cause a change in the correlation of the security's fundamental value with the fundamental values of other securities in the group.

One set of securities that satisfies these requirements is the S&P 500 index. Earlier we suggested identifying categories by looking at the products money managers offer their clients. The immense popularity of S&P-linked products suggests that this index may be a natural category in many investors' minds: S&P index funds and depositary receipts are important investment vehicles for both institutions and individuals, while S&P 500 futures are heavily traded by index arbitrageurs. The S&P 500 may also be a preferred habitat for U.S. investors who are reluctant to invest in foreign stocks and who doubt that active fund managers can outperform passive indices.

The S&P also has the second characteristic we require: there is clear and identifiable turnover in its membership. In a typical year there are about 30 changes; our full sample, which we describe in Section 3.1, includes 455 additions and 76 deletions.

Finally, the act of adding a stock to the S&P 500 should not cause a change in the covariance of the stock's cash flows with other stocks' cash flows. The stated goal of Standard and Poor's is to make the index representative of the U.S. economy, not to provide signals about future cash flows. Deletions from the index, however, are another matter. Stocks are usually removed from the index because a firm is merging, being taken over, or nearing bankruptcy. In these situations cash-flow characteristics may well be changing, so we exclude these cases from our deletion sample.

We therefore test Propositions 1 through 3 for the case where X is the S&P 500, and Y is stocks outside that index. In Section 3.2., in line with Proposition 1, we test whether a stock's beta with the S&P and the fraction of its variance explained by the index increase (decrease) after the stock's inclusion in (removal from) the index. In Section 3.3., in line with Proposition 2, we test whether a stock's beta with the S&P, controlling for the return of *non*-S&P stocks, goes up (falls) after inclusion (deletion). Finally, in Section 3.6., motivated by Proposition 3, we test whether the correlation of S&P and non-S&P stocks has fallen in line with the growing importance of the S&P as a category.

Our null hypothesis, laid out in Section 2.1., is that return comovement is primarily a function of comovement in news about fundamentals, so that the betas and R^2 just described, as well as the correlation of S&P and non-S&P stocks, do *not* change. The alternative hypothesis is that trading flows do induce comovement, and that the betas, R^2, and cross-category correlation change as predicted in the propositions.

While adding a stock to the S&P 500 should not *cause* a change in the cash-flow covariance matrix, it is possible that a stock's inclusion may *coincide* with a shift in the covariance matrix, and that this may drive some of our results. We address this possibility in Section 3.4.

3.1 Data

We consider S&P 500 index inclusions between September 22, 1976 and December 31, 2000 and deletions between January 1, 1979 and December 31, 2000. Standard & Poor's did not record announcement dates of index changes before September 1976 and we were unable to obtain data on deletions before 1979.

There are 590 inclusion events in the inclusion sample period and 565 deletions in the deletion sample period. Inclusion events are excluded if the new firm is a spin-off or a restructured version of a firm already in the index, if the firm is engaged in a merger or takeover around the inclusion event, or if required return data is not available. Deletion events are excluded if the firm is involved in a merger, takeover, or bankruptcy proceeding, or if required return data is not available.[11] These circumstances, determined by searching the NEXIS database, exclude the vast majority of deletions. The final sample includes 455 inclusions and 76 deletions.[12]

3.2 Univariate Regressions

If category-induced or habitat-induced trading flows cause return comovement, Proposition 1 predicts that stocks which are added to (deleted from) the S&P 500 will comove more (less) with the other members of the index after the addition or deletion event.

For each inclusion and deletion event, we run the univariate regression

$$R_{j,t} = \alpha_j + \beta_j R_{SP500,t} + v_{j,t} \tag{28}$$

separately for the period before the event and for the period after the event, and record the change in slope coefficient, $\Delta\beta_j$, and the change in R^2, ΔR^2_j. $R_{j,t}$ is the return of the stock involved in the change between time $t-1$ and t, while $R_{SP500,t}$ is the contemporaneous return on the S&P 500 index, obtained from the CRSP Index on the S&P Universe file.[13]

We run these regressions for three data frequencies: daily, weekly, and monthly. With daily and weekly data, the pre-event regression is run over the 12-month period ending the month before the month of the inclusion *announcement*, while the post-event regression is run over a 12-month period starting the month after the month of the inclusion *implementation*. In the case of monthly data, we use a 36-month period ending a month before the

[11] This last possibility may arise if the event occurs so close to the end of the sample that it prevents us from estimating post-event betas.

[12] The S&P 500 inclusion and deletion data are available upon request.

[13] In order to avoid spurious effects, we remove the contribution of the stock in question from the right-hand side variable. For addition events, this means that there are 500 stocks in the right-hand side variable before the addition, and 499 afterward. The reverse applies for deletion events.

announcement month and a 36-month period starting a month after the implementation month for the pre-event and post-event regressions, respectively.[14]

Table 1 reports the change in slope coefficient, averaged across all events in the sample, $\overline{\Delta\beta}$, as well as the average change in R^2, $\overline{\Delta R^2}$. It confirms that stocks added to the S&P 500 experience a strongly significant increase in daily and weekly betas and R^2. In the full sample of additions, the mean increase in daily beta is 0.151 and in weekly data, 0.11. At the monthly frequency, though, we are unable to detect a significant increase in either beta or R^2. Other than a weakly significant change in daily beta, we do not detect significant drops in beta or R^2 around deletion events.

Another prediction of trading-based comovement is that since the importance of the S&P as a category has grown over the course of our sample, the effects predicted by Proposition 1 should be stronger in the second half of our sample. Table 1 confirms that at daily and weekly frequencies, the increases in beta and R^2 across inclusion events are statistically stronger over the second subsample.

The standard errors in the table deserve comment. If two events are close together in calendar time, there may be substantial overlap in the time periods covered by the regressions associated with each event. This means that the disturbances $v_{j,t}$ may be correlated across events, which in turn implies that the $\Delta\beta_j$ may not be independent but rather autocorrelated at several lags.

We use simulation methods to compute standard errors that account for this dependence. We generate a simulated data set, consisting of an S&P return and returns on included stocks, and set the cross-sectional correlation of the disturbance terms to whatever value generates a first-order autocorrelation in the $\Delta\beta_j$'s equal to that observed in our results. We then compute $\overline{\Delta\beta}$ in this sample, under the null that betas do *not* change after inclusion. By generating many such data sets, we obtain the distribution of $\overline{\Delta\beta}$ under the null, and hence also, appropriate standard errors.[15]

[14]Up until October 1989, inclusions and deletions were made effective on the day of their announcement. Since then, the changes have been announced a few weeks in advance of their actual implementation. It is not clear whether to view the to-be-added stock as being in the index, or not in the index during the time between announcement and implementation; significant price effects have been documented on both days (Lynch and Mendenhall, 1997). To avoid these issues entirely, we do not use data from the month of the announcement or of the implementation; these are almost always the same month.

[15]It turns out that at least for daily and weekly frequencies, cross-correlation of disturbances does not affect the standard errors by very much. The reason is that such cross-correlation produces positive autocorrelation in the $\Delta\beta_j$ at the first few lags but negative autocorrelation at higher lags. As a result, the variance of $\overline{\Delta\beta}$ is only slightly higher than if the disturbances were uncorrelated.

3.3 Bivariate Regressions

The univariate regressions provide evidence of trading-based comovement at higher frequencies. Stronger evidence comes from tests of Proposition 2, which predicts that controlling for the return of *non*-S&P stocks, a stock that is added to or removed from the S&P will experience a large change in its loading on the S&P return. To test this, for each inclusion and deletion, we run the bivariate regression

$$R_{j,t} = \alpha_j + \beta_{j,SP500} R_{SP500,t} + \beta_{j,\text{non}SP500} R_{\text{non}SP500,t} + v_{j,t} \tag{29}$$

for the period before the event and the period after the event, and record the changes in S&P and non-S&P betas, $\Delta\beta_{j,SP500}$ and $\Delta\beta_{j,\text{non}SP500}$. $R_{\text{non}SP500,t}$ is the return on non-S&P stocks in the NYSE, AMEX, and Nasdaq universe between time $t-1$ and time t. This is inferred from index return and capitalization data using the identity that the capitalization-weighted average return of S&P stocks and of non-S&P stocks equals the overall CRSP value-weighted return on NYSE, AMEX, and Nasdaq stocks.

As before, we run the regressions at daily, weekly, and monthly frequencies. Daily and weekly regressions are run over a 12-month period ending the month before the announcement month and over a 12-month period starting the month after the implementation month. The monthly regressions use 36-month periods before announcement and after implementation.

Table 1 reports the change in S&P beta, averaged across all events in the sample, $\overline{\Delta\beta}_{SP500}$, as well as the average change in non-S&P beta, $\overline{\Delta\beta}_{\text{non}SP500}$. The results are stronger than the univariate results. At all three data frequencies, S&P 500 inclusion is associated with a substantial and significant increase in beta with the S&P and a substantial and significant decrease in beta with the rest of the market. For example, daily beta with the S&P 500 goes up by an average of 0.357 and daily beta with other stocks drops by -0.373. Large and significant results also obtain for deletion events at the daily and weekly frequencies. Moreover, the table confirms that at all three data frequencies, the changes in S&P and non-S&P betas are statistically stronger in the second subsample.

Figure I uses rolling regressions to show the dynamics of these changes. Panel A shows how the daily betas change over event time. The solid line shows the mean daily S&P beta and the dashed line shows the mean daily non-S&P beta. These coefficients are re-estimated each month using the prior 12 months of daily data. Therefore coefficients plotted to the left of the left vertical line use only pre-event returns. Coefficients plotted to the right of the right vertical line use only post-event returns. Coefficients in between use both pre- and post-event data. In terms of these figures, the beta changes reported in Table 1 are the average beta as of event month +12, which uses data from months [+1, +12] minus the average beta as of event month -1, which uses data from months [-12, -1]. There are fewer data points in the figures than in the table, however, because the figures include only firms with available return data for a full 24 months after inclusion. To be clear, the steady change in estimated betas between the two vertical lines should *not* be interpreted as a steady change in true betas. Rather, it arises from mixing data from the pre- and post-event regimes.

Our results on changes in S&P and non-S&P betas are consistent with the findings of Vijh (1994), who studies whether the rise of S&P-linked products has affected the standard measure of stock risk, namely beta with respect to the *overall* market return. He finds that over the 1975-1989 period, a stock's daily beta with the market goes up by a statistically significant 0.08, on average, after inclusion. Since a large fraction of overall market value comes from S&P stocks, this fits with the increase in S&P beta we detect over a similar time period. Given our result that non-S&P beta falls significantly, it also makes sense that the rise in overall market beta should be considerably smaller than the rise in S&P beta.[16]

3.4 Evaluating Alternative Explanations

We now consider two alternative explanations for the results in Table 1 and Figure I. One possibility is that stocks in the S&P 500 index differ from other stocks in terms of some characteristic, and that the stocks Standard and Poor's chooses to include are stocks that are increasingly demonstrating that characteristic. If the characteristic is also associated with a cash-flow factor, this may explain our results.

The most obvious such characteristic is size. Stocks in the S&P have considerably higher market capitalizations than stocks outside the index, and the stocks Standard and Poor's includes into the index have often been growing in size prior to inclusion. Moreover, size is associated with a cash-flow factor: there is a common component to news about the earnings of large-cap stocks. Our finding that S&P betas increase around inclusion may simply reflect the fact that included stocks are growing in size around inclusion and are therefore increasingly loading on the large stock cash-flow factor. More generally, this is a story in which inclusion into the S&P *coincides* with a change in the cash-flow covariance matrix, even if it does not cause it.

Another potential explanation is based on industry effects. Suppose that some industry becomes increasingly dominant in the economy. This increases the fraction of the value of the S&P made up by stocks in this industry. Moreover, in an effort to keep their index representative, Standard and Poor's may start drawing an increasing number of new inclusions from this industry. Since S&P beta is computed using the *value*-weighted S&P return,

[16]Under the CAPM, an increase in overall market beta after inclusion predicts that stocks should drop in price when they are added to the index. The fact that such stocks actually display large price *increases* upon inclusion clearly rejects this prediction.

this simultaneity can in principle explain our results: if Yahoo! is included into the S&P at precisely the time that other technology stocks in the index are growing in value – as indeed it was, having been added in December 1999 – it may covary more with the S&P after inclusion than before.

To address both these competing explanations, we perform a matching exercise. For each event stock included into the S&P during our sample period, we search for a "matching" stock, drawn from the same industry as the event stock and in the same size decile as the event stock, both at the time of inclusion and 12 months before inclusion, *but which is not included into the index*. In other words, since the matching stock matches the event stock on industry and on recent growth in market capitalization, it is as good a candidate for inclusion as the event stock itself, but simply happens not to be included. If the matching stocks do not demonstrate the same increase (decrease) in S&P (non-S&P) betas as the event stocks, it strengthens the case that the results in Table 1 and Figure I are due to trading-based comovement, rather than to the alternative explanations.[17] In the case of deleted stocks, the matching stock is a stock in the S&P which matches the deleted stock on industry, and recent change in market capitalization, but which is not removed from the index.

Table 2 and Figure II contain the results of the matching exercise. Figure II, constructed in a parallel fashion to Figure I, presents the evolution of S&P and non-S&P betas in event time for *matching* stocks. Table 2 reports the change in betas and R^2 in univariate and bivariate regressions for event stocks relative to the analogous changes for matching stocks.

Figure II suggests that at daily and weekly frequencies, the alternative stories can explain only a small fraction of our results: the matching stocks exhibit much smaller shifts in betas than do the event stocks. However, it also suggests that at the monthly frequency, the characteristic- and industry-based explanations *do* have some bite: even for matching stocks, S&P betas increase around inclusion and non-S&P betas decrease.

Table 2 confirms these impressions. At the daily and weekly frequency, the changes in beta and R^2 in univariate regressions and in S&P and non-S&P betas in bivariate regressions, remain strongly significant across inclusion events, even after subtracting off the corresponding changes for matching stocks. At the monthly level, however, a substantial part of the strong bivariate regression results in Table 1 are explained by the matching stocks. Nonetheless, in the second subsample, the increase in S&P beta for event stocks is still significantly

[17] At the monthly frequency, in order to match the window betas are computed over, we look for matching stocks that match the event stock on size both at inclusion and 36 months before inclusion. At all frequencies, we initially try to match by SIC4 industry code. If no match can be found, we allow the matching stock to be in the same SIC3 industry class, then to be within one size decile at inclusion, then to be within one size decile 12 months before inclusion, then to be in the same SIC2 industry class, then to be within two size deciles at inclusion, then to be within two size deciles 12 months before inclusion, and finally to be within three size deciles 12 months before inclusion. Events for which no such matches can be found are not included in the matching exercise samples.

greater than that for matching stocks.[18]

Overall, it appears that trading-based comovement operates strongly at daily and weekly frequencies; at the monthly frequency, its effects are still present, but are less pronounced.

3.5 Calendar Time Tests

The methodology we use to test Propositions 1 and 2 in Sections 3.2 and 3.3 is often called an "event time" approach. An alternative methodology is a "calendar time" approach. This technique is often used to address a common statistical problem in event studies, namely correlation of returns across events. As described in Section 3.2., we use simulations to deal with this issue. Performing calendar time tests offers a second way of checking that our results are robust to these statistical considerations.

The calendar time approach requires the construction of two portfolios: a "pre-event" portfolio whose return at time t, $R_{pre,t}$, is the equal-weighted average return at time t of all stocks that *will* be added to the index within some window after time t; and a "post-event" portfolio whose return at time t, $R_{post,t}$, is the equal-weighted average return at time t of all stocks that *have been* added to the index within some window preceding time t. In our analyses of daily and weekly data, we take the window to be a year, and extend it to three years for monthly data.

The calendar time test of Proposition 1 then calls for running two regressions:

$$R_{pre,t} = \alpha_{pre} + \beta_{pre} R_{SP500,t} + v_{pre,t} \tag{30}$$

and

$$R_{post,t} = \alpha_{post} + \beta_{post} R_{SP500,t} + v_{post,t} \tag{31}$$

and checking whether $\beta_{post} > \beta_{pre}$ and whether the R^2 in the second regression is greater than in the first.

Similarly, the calendar time test of Proposition 2 calls for running the following two regressions,

$$R_{pre,t} = \alpha_{pre} + \beta_{pre,SP500} R_{SP500,t} + \beta_{pre,\mathrm{non}SP500} R_{\mathrm{non}SP500,t} + v_{pre,t} \tag{32}$$

and

$$R_{post,t} = \alpha_{post} + \beta_{post,SP500} R_{SP500,t} + \beta_{post,\mathrm{non}SP500} R_{\mathrm{non}SP500,t} + v_{post,t} \tag{33}$$

[18] In Table 1, we conducted simulations to correct the standard errors for possible correlation in disturbance terms across regressions. This problem affects matching stock regressions just as much as it does event stock regressions, but it does not affect *differences* in slopes across the two sets of regressions. The Table 2 standard errors are therefore the usual ones – no simulation-based correction is required.

and checking whether $\beta_{post,SP500} > \beta_{pre,SP500}$ and $\beta_{post,\text{non}SP500} < \beta_{pre,\text{non}SP500}$.

Table 3 reports the changes in slope coefficients and R^2s. In general, the results are as supportive of trading-based comovement as the event time tests. In the univariate regressions, significant increases in beta and R^2 occur at the daily and weekly frequencies, and for R^2, even at the monthly frequency. In the bivariate regressions, the results for inclusion events are strongly significant at all three data frequencies, although the results for deletion events are weaker than before: there is no statistically significant effect at any frequency.

3.6 Comovement Across Categories

Proposition 3 predicts that the correlation of the returns of two groups of securities will be lower than the correlation of their fundamentals if these groups form natural categories or habitats. This proposition is testable in the time series under the condition that the groups' importance as categories or habitats has grown over time.

The S&P 500 satisfies this last condition: its use in various investment styles has grown dramatically in the last few decades. Consistent with this trend, Wurgler and Zhuravskaya (2001) find that the size of the inclusion price jump has grown with the volume of funds devoted to S&P indexing, and our earlier results show increasing comovement effects in more recent years.

Table 4 reports the trends in comovement between the S&P and other stocks over the past thirty years. The left column shows that the relative size of the S&P and whole market has remained constant. The declining correlations in the right columns show that at all three data frequencies, the returns on the S&P 500 have grown increasingly divorced from the returns on the rest of the market. The correlation in returns remains high today, but it is not as high as it was prior to the advent of the S&P 500 as a category. Another interesting pattern is that the decline in the daily correlation seems to have halted in recent years, while the weekly and monthly correlations continue to decline.

In Table 5 we determine whether the decreasing correlation between S&P and non-S&P stocks is statistically significant, or whether the correlation between two random groups would on average display a similar decline. We construct value-weighted returns on a random group of 500 stocks and compute their correlation with the value-weighted returns on the rest of the market over consecutive five year periods. By repeating this procedure for many random groups of 500 stocks, we can construct sampling distributions for the change in correlation over various intervals. We can then determine whether the decline in the S&P correlation is unusually large.

The left columns of Table 5 report the sampling distribution of the changes in correlation between the random 500 and the rest of the market. The correlations between random groups of stocks have declined. Panel A shows that, from the early 1970s to the late 1990s, the daily return correlation between random groups has fallen by a median of -0.043. For comparison, the second column from the right reports the experience of the S&P 500. Over this same period, Table 4 indicates that the daily return correlation between the S&P and the rest of the market has fallen by -0.118. The last column indicates that this is a much greater decline than expected by chance. A similar conclusion emerges for weekly data. At the monthly level, the decline in correlation between S&P and non-S&P stocks is below the average decline for randomly-chosen stocks, but is not statistically unusual.

Our simulation controls for the possibility that the decline in the S&P and non-S&P return correlation is due to a general decline in the correlation of stock fundamentals. Indeed, the results of Campbell, Lettau, Malkiel, and Xu (2001) suggest that such a decline in fundamental correlation has occured, making it important to control for. Our simulation does not, however, rule out the possibility that our results are due to an especially large decline in the correlation of S&P 500 stocks' fundamentals with remaining stocks' fundamentals, as compared to the decline in the correlation of a random 500 stocks' fundamentals with remaining stocks' fundamentals. However, we see no obvious reason why this would be the case, since the S&P 500 index has always been constructed to be representative of the overall economy.[19]

4 Conclusion

In this paper, we present and examine empirically three models of comovement. The traditional model attributes comovement to correlation in news about fundamental value. The two alternative models we consider explain comovement by correlated investor demand shifts for securities in a given category, or by demand shifts by specific investor clienteles.

To assess these theories, we consider the well-studied phenomenon of stock inclusions into, and deletions from, the S&P 500 index. While previous studies have noted significant immediate price effects associated with inclusions and deletions, we focus on changes in the patterns of comovement of newly included (or deleted) stocks with stocks already in the index. We find that stocks included into the index begin to comove more with other stocks in the index, and less with stocks out of the index. The converse holds for deletions. Because inclusion into the S&P 500 index conveys no news about fundamentals, this evidence is hard

[19]Panel A of Table 4 also shows that the abrupt halt in the decline of the daily S&P correlation after 1990 is not mirrored by the random-500 correlation, while the weekly and monthly S&P correlations continue to decline relative to the typical random-500 group. One explanation is that arbitrage has checked the decline in the daily correlation, but has yet to stop the decline in the weekly and monthly correlations. De Long et al. (1990) point out that long-horizon arbitrage is likely to be weaker than short-horizon arbitrage.

to reconcile with the fundamentals view of comovement, but supports the theories based on shifts in demand.

This evidence adds to the growing range of phenomena identified by financial economists that reveal the importance of asset classification, and of demand shifts among asset classes, for valuation. From this perspective, a security's price may depend not only on its fundamentals, but also on which asset categories it belongs to, and on which investors trade it.

5 Appendix

Proof of Propositions 1, 2, and 3: Suppose that asset $n+1$ is reclassified from Y into X, and that at the same moment, asset 1 is reclassified from X into Y. Before reclassification,

$$\begin{aligned}
\Delta P_{X,t+1} &= \varepsilon_{X,t+1} + \frac{\Delta u_{X,t+1}}{\phi_1} + \frac{\Delta u_{Y,t+1}}{\phi_2} \\
\Delta P_{Y,t+1} &= \varepsilon_{Y,t+1} + \frac{\Delta u_{X,t+1}}{\phi_2} + \frac{\Delta u_{Y,t+1}}{\phi_1} \\
\Delta P_{n+1,t+1} &= \varepsilon_{n+1,t+1} + \frac{\Delta u_{X,t+1}}{\phi_2} + \frac{\Delta u_{Y,t+1}}{\phi_1},
\end{aligned} \tag{34}$$

where

$$\varepsilon_{k,t} = \frac{1}{n} \sum_{l \epsilon k} \varepsilon_{l,t},\ k = X, Y.$$

This implies, as $n \to \infty$,

$$\begin{aligned}
\text{cov}(\Delta P_{n+1,t+1}, \Delta P_{X,t+1}) &= \psi_M^2 + \frac{2\sigma_u^2}{\phi_1\phi_2} + \sigma_u^2 \rho_u (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}) \\
\text{cov}(\Delta P_{n+1,t+1}, \Delta P_{Y,t+1}) &= \psi_M^2 + \psi_S^2 + \sigma_u^2 (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}) + \frac{2\sigma_u^2 \rho_u}{\phi_1\phi_2} \\
\text{var}(\Delta P_{X,t+1}) &= \text{var}(\Delta P_{Y,t+1}) = \psi_M^2 + \psi_S^2 + \sigma_u^2 (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}) + \frac{2\sigma_u^2 \rho_u}{\phi_1\phi_2} \\
\text{cov}(\Delta P_{X,t+1}, \Delta P_{Y,t+1}) &= \psi_M^2 + \frac{2\sigma_u^2}{\phi_1\phi_2} + \sigma_u^2 \rho_u (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}).
\end{aligned} \tag{35}$$

After reclassification, $\Delta P_{X,t+1}$ and $\Delta P_{Y,t+1}$ are still given by (34), but now

$$\Delta P_{n+1,t+1} = \varepsilon_{n+1,t+1} + \frac{\Delta u_{X,t+1}}{\phi_1} + \frac{\Delta u_{Y,t+1}}{\phi_2}. \tag{36}$$

This implies, as $n \to \infty$,

$$\begin{aligned}
\text{cov}\ (\Delta P_{n+1,t+1}, \Delta P_{X,t+1}) &= \psi_M^2 + \sigma_u^2 (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}) + \frac{2\sigma_u^2 \rho_u}{\phi_1\phi_2}, \\
\text{cov}\ (\Delta P_{n+1,t+1}, \Delta P_{Y,t+1}) &= \psi_M^2 + \psi_S^2 + \frac{2\sigma_u^2}{\phi_1\phi_2} + \sigma_u^2 \rho_u (\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2}),
\end{aligned} \tag{37}$$

while $\text{var}(\Delta P_{X,t})$, $\text{var}(\Delta P_{Y,t})$, and $\text{cov}(\Delta P_{X,t+1}, \Delta P_{Y,t+1})$ remain the same as before.

Since the OLS estimate of β_{n+1} in the regression

$$\Delta P_{n+1,t+1} = \alpha_{n+1} + \beta_{n+1} \Delta P_{X,t+1} + v_{n+1,t+1} \tag{38}$$

is given by

$$\beta_{n+1} = \frac{\text{cov}(\Delta P_{n+1,t+1}, \Delta P_{X,t+1})}{\text{var}(\Delta P_{X,t+1})}, \tag{39}$$

expressions (35) and (37) taken together with

$$\frac{1}{\phi_1^2} + \frac{1}{\phi_2^2} - \frac{2}{\phi_1\phi_2} = (\frac{1}{\phi_1} - \frac{1}{\phi_2})^2 \geq 0,$$

confirm that β_{n+1} increases after reclassification as claimed in Proposition 1. Moreover, since $\mathrm{var}(\Delta P_{n+1,t})$ and $\mathrm{var}(\Delta P_{X,t})$ are unchanged after reclassification, the increase in β_{n+1} also implies an increase in R^2 of regression (38) after inclusion.

The OLS estimates of $\beta_{n+1,X}$ and $\beta_{n+1,Y}$ in the regression

$$\Delta P_{n+1,t+1} = \alpha_{n+1} + \beta_{n+1,X}\Delta P_{X,t+1} + \beta_{n+1,Y}\Delta P_{Y,t+1} + v_{n+1,t+1} \tag{40}$$

are given by

$$\begin{pmatrix} \beta_{n+1,X} \\ \beta_{n+1,Y} \end{pmatrix} = \frac{1}{V_X V_Y - C_{XY}^2} \begin{pmatrix} V_Y & -C_{XY} \\ -C_{XY} & V_X \end{pmatrix} \begin{pmatrix} C_{n+1,X} \\ C_{n+1,Y} \end{pmatrix} \tag{41}$$

where

$$\begin{aligned} V_k &= \mathrm{var}(\Delta P_{k,t+1}),\ k = X, Y \\ C_{XY} &= \mathrm{cov}(\Delta P_{X,t+1}, \Delta P_{Y,t+1}) \\ C_{n+1,k} &= \mathrm{cov}(\Delta P_{n+1,t+1}, \Delta P_{k,t+1}),\ k = X, Y. \end{aligned}$$

Before reclassification, $C_{n+1,X} = C_{XY}$ and $C_{n+1,Y} = V_Y$, while after reclassification, $C_{n+1,X} = V_X - \psi_S^2$ and $C_{n+1,Y} = C_{XY} + \psi_S^2$. It is easy to check that this implies that $\beta_{n+1,X}$ does indeed increase after reclassification, while $\beta_{n+1,Y}$ falls. This proves Proposition 2.

Finally, given the expressions for $\mathrm{var}(\Delta P_{X,t+1})$, $\mathrm{var}(\Delta P_{Y,t+1})$, and $\mathrm{cov}(\Delta P_{X,t+1}, \Delta P_{Y,t+1})$ in equation (35), it is immediate that

$$\mathrm{corr}(\Delta P_{X,t+1}, \Delta P_{Y,t+1}) < \mathrm{corr}(\Delta D_{X,t+1}, \Delta D_{Y,t+1}).$$

This proves Proposition 3.

6 References

Barberis, N., and A. Shleifer (2003), "Style Investing," forthcoming, *Journal of Financial Economics.*

Bodurtha J., Kim, D., and C.M. Lee (1995), "Closed-end Country Funds and U.S. Market Sentiment," *Review of Financial Studies* 8, 879-918.

Campbell, J.Y., Lettau, M., Malkiel, B., and Y. Xu (2001), "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance* 56, 1-43.

De Long, J.B., Shleifer A., Summers L., and R. Waldmann (1990), "Noise Trader Risk in Financial Markets," *Journal of Political Economy* 98, 703-38.

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

Fama, E., and K. French (1995), "Size and Book-to-Market Factors in Earnings and Returns," *Journal of Finance* 50, 131-155.

French, K., and R. Roll (1986), "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics* 17, 5-26.

Froot, K., and E. Dabora (1999), "How are Stock Prices affected by the Location of Trade?," *Journal of Financial Economics* 53, 189-216.

Goetzmann, W., and M. Massa (2001), "Index funds and stock market growth," Working Paper, Yale University.

Gompers, P., and A. Metrick (2001), "Institutional Investors and Equity Prices," *Quarterly Journal of Economics* 116, 229-259.

Greenwood, R. (2001), "Large Events and Limited Arbitrage: Evidence from a Japanese Stock Index Redefinition," Working Paper, Harvard University.

Greenwood, R., and N. Sosner (2002), "Where Do Betas Come From?," Working Paper, Harvard University.

Hardouvelis, G., La Porta R., and T. Wizman (1994), "What Moves the Discount on Country Equity Funds?," in Jeffrey Frankel (ed.), *The Internationalization of Equity Markets*, Chicago, The University of Chicago Press.

Harris, L., and E. Gurel (1986), "Price and Volume Effects Associated with Changes in the S&P 500: New Evidence for the Existence of Price Pressure," *Journal of Finance* 41, 851-860.

Kaul, A., Mehrotra V., and R. Morck (2000), "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment," *Journal of Finance* 55, 893-912.

Lamont, O., and R. Thaler (2000), "Can the Market Add and Subtract? Mispricing in Tech Stock Carveouts," Working Paper, University of Chicago.

Lee, C., Shleifer A., and R. Thaler (1991), "Investor Sentiment and the Closed-end Fund puzzle," *Journal of Finance* 46, 75-110.

Lynch, A. and R. Mendenhall (1997), "New evidence on stock price effects associated with changes in the S&P 500 Index," *Journal of Business* 70, 351-83.

Merton, R. (1987), "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance* 42, 483-510.

Mitchell, M., Pulvino, T., and E. Stafford (2002), "Limited Arbitrage in Equity Markets," forthcoming, *Journal of Finance*.

Pindyck, R., and J. Rotemberg (1990), "The Excess Comovement of Commodity Prices," *Economic Journal* 100, 1173-1189.

Pindyck, R. and J. Rotemberg (1993), "The Comovement of Stock Prices," *Quarterly Journal of Economics* 108, 1073-1104.

Shiller, R. (1989), "Comovements in Stock Prices and Comovements in Dividends," *Journal of Finance* 46, 719-729.

Shleifer A. (1986), "Do Demand Curves for Stocks Slope Down?," *Journal of Finance* 41, 579-590.

Shleifer A., and R. Vishny (1997), "The Limits of Arbitrage," *Journal of Finance* 52, 35-55.

Vijh, A. (1994), "S&P 500 Trading Strategies and Stock Betas," *Review of Financial Studies* 7, 215-251.

Wurgler, J., and K. Zhuravskaya (2001), "Does Arbitrage Flatten Demand Curves for Stocks?," forthcoming, *Journal of Business*.

Table 1. Changes in comovement of stocks added to and deleted from the S&P 500 Index. Changes in the slope and the fit of regressions of returns of stocks added to and deleted from the S&P 500 Index on returns of the S&P 500 Index and the non-S&P 500 rest of the market. The sample includes stocks added to and deleted from the S&P 500 between 1976 and 2000 which were not involved in mergers or related events (described in the text), and which have sufficient return data on CRSP. For each added or deleted stock *j*, the univariate model

$$R_{j,t} = a_j + b_j R_{SP500,t} + u_{j,t}$$

and the bivariate model

$$R_{j,t} = a_j + b_{j,SP500} R_{SP500,t} + b_{j,nonSP500} R_{nonSP500,t} + u_{j,t}$$

are separately estimated for the pre-change and post-change period. Returns on the S&P 500 (R_{SP500}) are from the CRSP Index on the S&P 500 Universe file. Returns on a capitalization-weighted index of the non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq are inferred from the identity

$$R_{VWCRSP,t} = \left(\frac{CAP_{CRSP,t-1} - CAP_{SP500,t-1}}{CAP_{CRSP,t-1}}\right) R_{nonSP500,t} + \left(\frac{CAP_{SP500,t-1}}{CAP_{CRSP,t-1}}\right) R_{SP500,t} .$$

Total capitalization on the S&P 500 (CAP_{SP500}) is from the CRSP Index on the S&P 500 Universe file. Returns on the value-weighted CRSP NYSE, AMEX, and Nasdaq index (R_{VWCRSP}) and total capitalization (CAP_{CRSP}) are from the CRSP Stock Index file. Returns from October 1987 are excluded. The mechanical influence of the added or deleted stock is removed from the independent variables as appropriate. For the univariate regression model, we examine the mean difference between the pre-change slope and the post-change slope $\overline{\Delta b}$, and the mean change in fit $\overline{\Delta R^2}$. For the bivariate model, we examine the mean changes in the slopes, $\overline{\Delta b}_{SP500}$ and $\overline{\Delta b}_{nonSP500}$. The pre-change and post-change estimation periods are [-12,-1] and [+1,+12] months for daily and weekly returns and [-36,-1] and [+1,+36] months for monthly returns. Panels A, B, and C show results for daily, weekly, and monthly returns, respectively. Standard errors are determined by simulation, to account for cross-correlation, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels in one-sided tests, respectively.

Sample		N	Univariate		Bivariate	
			$\overline{\Delta b}$ (s.e.)	$\overline{\Delta R^2}$ (s.e.)	$\overline{\Delta b_{SP500}}$ (s.e.)	$\overline{\Delta b_{nonSP500}}$ (s.e.)
Panel A. Daily Returns						
Additions	1976-2000	455	0.151*** (0.021)	0.049*** (0.005)	0.357*** (0.022)	-0.373*** (0.029)
	1976-1987	196	0.067*** (0.023)	0. 038*** (0.008)	0.252*** (0.041)	-0.262*** (0.050)
	1988-2000	259	0.214*** (0.032)	0.058*** (0.007)	0.406*** (0.027)	-0.426*** (0.035)
Deletions	1976-2000	76	-0.087* (0.049)	-0.010 (0.007)	-0.511*** (0.111)	0.550*** (0.122)
Panel B. Weekly Returns						
Additions	1976-2000	455	0.110*** (0.029)	0.033*** (0.008)	0.174*** (0.053)	-0.119** (0.056)
	1976-1987	196	0.025 (0.036)	0.027** (0.012)	0.137 (0.094)	-0.125 (0.093)
	1988-2000	259	0.173*** (0.043)	0.037*** (0.010)	0.202*** (0.061)	-0.115* (0.069)
Deletions	1976-2000	76	-0.129 (0.105)	-0.015 (0.010)	-0.505*** (0.161)	0.412** (0.169)
Panel C. Monthly Returns						
Additions	1976-1998	324	0.042 (0.041)	0.004 (0.014)	0.317*** (0.077)	-0.252*** (0.072)
	1976-1987	172	-0.010 (0.060)	0.006 (0.021)	0.267** (0.127)	-0.167 (0.116)
	1988-1998	152	0.101 (0.066)	0.000 (0.021)	0.375*** (0.113)	-0.348*** (0.107)
Deletions	1976-1998	45	0.006 (0.100)	0.001 (0.022)	0.303 (0.240)	-0.256 (0.252)

Table 2. Changes in comovement of stocks added to and deleted from the S&P 500 Index: Relative to matching firms. Changes in the slope and the fit of regressions of returns on stocks added to and deleted from the S&P 500 Index relative to changes in the same parameters for matching stocks. Each stock in the event sample is paired with another stock which matches it on industry, market capitalization and growth in market capitalization over the pre-change estimation period (described in text). The event sample includes stocks added to and deleted from the S&P 500 between 1976 and 2000 which were not involved in mergers or related events, which have sufficient return data on CRSP, and for which a matching stock could be found. For each added or deleted stock *j*, the univariate model

$$R_{j,t} = a_j + b_j R_{SP500,t} + u_{j,t}$$

and the bivariate model

$$R_{j,t} = a_j + b_{j,SP500} R_{SP500,t} + b_{j,nonSP500} R_{nonSP500,t} + u_{j,t}$$

are separately estimated for the pre-change and post-change period, and analogous regressions are run for each matching stock. Returns on the S&P 500 (R_{SP500}) are from the CRSP Index on the S&P 500 Universe file. Returns on a capitalization-weighted index of the non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq are inferred from the identity described in Table 1. Returns from October 1987 are excluded. The mechanical influence of the added or deleted stock is removed from the independent variables as appropriate. For the univariate regression model, we examine the mean difference between the pre- and post-change slope and fit of the event stock and the matching stock, $\overline{\Delta\Delta b}$ and $\overline{\Delta\Delta R^2}$. For the bivariate model, we examine the mean difference between the changes in the slopes of the event stock and the matching stock, $\overline{\Delta\Delta b_{SP500}}$ and $\overline{\Delta\Delta b_{nonSP500}}$. The pre-change and post-change estimation periods are [-12,-1] and [+1,+12] months for daily and weekly returns and [-36,-1] and [+1,+36] months for monthly returns. Panels A, B, and C show results for daily, weekly, and monthly returns, respectively. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels in one-sided tests, respectively.

Sample		N	Univariate $\overline{\Delta\Delta b}$ (s.e.)	Univariate $\overline{\Delta\Delta R^2}$ (s.e.)	Bivariate $\overline{\Delta\Delta b_{SP500}}$ (s.e.)	Bivariate $\overline{\Delta\Delta b_{nonSP500}}$ (s.e.)
Panel A. Daily Returns						
Additions	1976-2000	435	0.120***	0.040***	0.318***	-0.289***
			(0.022)	(0.006)	(0.035)	(0.042)
	1976-1987	189	0.109***	0. 033***	0.262***	-0.257***
			(0.028)	(0.008)	(0.051)	(0.065)
	1988-2000	246	0.129***	0.046***	0.361***	-0.313***
			(0.032)	(0.007)	(0.047)	(0.055)
Deletions	1976-2000	36	-0.098	-0.012	-0.298**	0.271
			(0.081)	(0.013)	(0.142)	(0.193)
Panel B. Weekly Returns						
Additions	1976-2000	434	0.077**	0.028***	0.208***	-0.146**
			(0.037)	(0.009)	(0.070)	(0.070)
	1976-1987	188	0.086*	0.026*	0.202*	-0.162
			(0.047)	(0.015)	(0.113)	(0.114)
	1988-2000	246	0.070	0.030***	0.212**	-0.134
			(0.055)	(0.012)	(0.087)	(0.088)
Deletions	1976-2000	36	-0.013	-0.030	-0.616**	0.771**
			(0.157)	(0.020)	(0.251)	(0.285)
Panel C. Monthly Returns						
Additions	1976-1998	300	0.104**	0.011	0.173*	-0.090
			(0.047)	(0.013)	(0.103)	(0.091)
	1976-1987	162	0.054	0.008	0.054	-0.011
			(0.060)	(0.020)	(0.145)	(0.125)
	1988-1998	138	0.163**	0.015	0.313**	-0.183
			(0.073)	(0.020)	(0.144)	(0.133)
Deletions	1976-1998	18	0.236	0.057	0.438	-0.133
			(0.156)	(0.041)	(0.315)	(0.266)

Table 3. Changes in comovement of stocks added to and deleted from the S&P 500 Index: Calendar time. Differences between the comovement characteristics of two portfolios of stocks: those about to be added to the S&P 500 and those just recently added. The sample includes stocks added to and deleted from the S&P 500 between 1976 and 2000 which were not involved in mergers or related events, and which have sufficient return data on CRSP. A capitalization-weighted return index of non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq is inferred from the identity described in Table 1. Returns from October 1987 are excluded. In daily data, for example, each day we form an equal-weighted portfolio of stocks that will be added to the S&P 500 within the next year and a portfolio of stocks that were added within in the past year. We then run separate univariate regressions for each portfolio on the S&P 500 index,

$$R_{pre,t} = a_{pre} + b_{pre} R_{SP500,t} + u_{pre,t} \quad \text{and}$$
$$R_{post,t} = a_{post} + b_{post} R_{SP500,t} + u_{post,t} \, ,$$

denoting the difference in slope and fit between the "post" and "pre" regressions as Δb and ΔR^2, respectively. We also run separate bivariate regressions for each portfolio,

$$R_{pre,t} = a_{pre} + b_{pre,SP500} R_{SP500,t} + b_{pre,nonSP500} R_{nonSP500,t} + u_{pre,t} \quad \text{and}$$
$$R_{post,t} = a_{post} + b_{post,SP500} R_{SP500,t} + b_{post,nonSP500} R_{nonSP500,t} + u_{post,t} \, ,$$

denoting the difference in the slopes as Δb_{SP500} and $\Delta b_{nonSP500}$, respectively. The mechanical influence of the pre and post portfolio stocks is removed, as appropriate, from the independent variables. In daily and weekly data, the pre portfolio includes stocks that will be added within one year and the post portfolio includes stocks that were added in the past year. In monthly data, these windows are extended to three years. We require at least 10 stocks in each portfolio in order for that observation (day, month, or year) to be included in the regressions. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels in one-sided tests, respectively.

Sample		T	Univariate Δb (s.e.)	Univariate ΔR^2 (s.e.)	Bivariate Δb_{SP500} (s.e.)	Bivariate $\Delta b_{nonSP500}$ (s.e.)
Panel A. Daily Returns						
Additions	1976-2000	4147	0.123*** (0.013)	0.100*** (0.007)	0.297*** (0.020)	-0.262*** (0.022)
	1976-1987	1873	0.116*** (0.019)	0.061*** (0.010)	0.326*** (0.035)	-0.329*** (0.044)
	1988-2000	2274	0.129*** (0.017)	0.119*** (0.010)	0.298*** (0.026)	-0.247*** (0.026)
Deletions	1976-2000	151	0.058 (0.107)	-0.069 (0.063)	0.164 (0.190)	-0.080 (0.117)
Panel B. Weekly Returns						
Additions	1976-2000	856	0.045* (0.026)	0.028* (0.015)	0.167*** (0.046)	-0.154*** (0.043)
	1976-1987	387	0.041 (0.039)	-0.028 (0.024)	-0.014 (0.094)	0.051 (0.099)
	1988-2000	469	0.049 (0.036)	0.065*** (0.022)	0.219*** (0.054)	-0.210*** (0.048)
Deletions	1976-2000	29	-0.082 (0.193)	-0.160 (0.120)	-0.162 (0.537)	0.039 (0.288)
Panel C. Monthly Returns						
Additions	1976-2000	282	0.018 (0.042)	0.090*** (0.030)	0.319*** (0.073)	-0.320*** (0.061)
	1976-1987	127	0.019 (0.047)	0.007 (0.048)	0.148 (0.114)	-0.143 (0.099)
	1988-2000	155	0.016 (0.065)	0.157** (0.067)	0.388*** (0.092)	-0.406*** (0.075)
Deletions	1976-2000	116	-0.123 (0.105)	-0.293*** (0.058)	-0.255 (0.166)	0.132 (0.132)

Table 4. Trends in the correlation between returns on the S&P 500 and the rest of the market. The correlation between the S&P 500 and the rest of the market. Returns on the S&P 500 (R_{SP500}) are from the CRSP Index on the S&P 500 Universe file. Returns on a capitalization-weighted index of the non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq are inferred from the identity

$$R_{VWCRSP,t} = \left(\frac{CAP_{CRSP,t-1} - CAP_{SP500,t-1}}{CAP_{CRSP,t-1}}\right) R_{nonSP500,t} + \left(\frac{CAP_{SP500,t-1}}{CAP_{CRSP,t-1}}\right) R_{SP500,t}.$$

Total capitalization the S&P 500 (CAP_{SP500}) is from the CRSP Index on the S&P 500 Universe file. Returns on the value-weighted CRSP NYSE, AMEX, and Nasdaq index (R_{VWCRSP}) and total capitalization (CAP_{CRSP}) are from the CRSP Stock Index file. Returns from October 1987 are excluded.

Years	Mean of $\left(\frac{CAP_{SP500,t}}{CAP_{CRSP,t}}\right)$	Correlation between S&P 500 and the rest of the market		
		Daily Returns	Weekly Returns	Monthly Returns
1970 – 1974	0.689	0.941	0.942	0.931
1975 – 1979	0.685	0.898	0.920	0.925
1980 – 1984	0.670	0.871	0.915	0.929
1985 – 1989	0.683	0.825	0.864	0.912
1990 – 1994	0.690	0.817	0.851	0.901
1995 – 1999	0.701	0.823	0.810	0.840

Table 5. Trends in the correlation between returns on the S&P 500 and the rest of the market: Relative to random 500 stocks. The decline in correlation between returns on the S&P 500 and the rest of the market is evaluated relative to the null hypothesis that a similar decline applies to random categories of stocks. The distribution of changes in the correlation between the return on 500 random stocks from NYSE, AMEX, and Nasdaq and the value-weighted return on the rest of the market is determined by simulation. The following procedure is repeated 500 times: (i) A sample of 500 random stocks from the NYSE, AMEX, and Nasdaq is identified from all stocks that CRSP lists for 1970. The complementary set of stocks, i.e. the rest of the market, is also identified as of 1970. (ii) The daily, weekly, and monthly correlation between these two portfolios is computed and recorded each year from 1970 through 1999. If a stock drops out of the random 500 sample, it is replaced with a stock randomly taken from the rest of the market sample. Returns from October 1987 are excluded. (iii) These two return series represent one sample path, over which correlations and changes in correlations can be estimated. Panels A, B, and C show results for daily, weekly, and monthly returns, respectively.

		Change in correlation between random 500 and the rest of the market					Change in correlation between S&P 500 and the rest of the market	
Years		5th percentile	10^{th} percentile	50^{th} percentile	90^{th} percentile	95^{th} percentile	Actual	Prob (Actual < X)
				Panel A. Daily Returns				
1995 – 1999 vs.	1970 – 1974	- 0.070	- 0.064	-0.043	- 0.023	- 0.019	- 0.118	0.002
	1975 – 1979	- 0.073	- 0.066	-0.045	- 0.028	- 0.024	- 0.075	0.040
	1980 – 1984	- 0.070	- 0.065	-0.042	- 0.023	- 0.019	- 0.048	0.344
	1985 – 1989	- 0.079	- 0.073	-0.053	- 0.037	- 0.032	- 0.002	0.996
	1990 – 1994	- 0.048	- 0.042	-0.026	- 0.012	- 0.009	+ 0.006	0.998

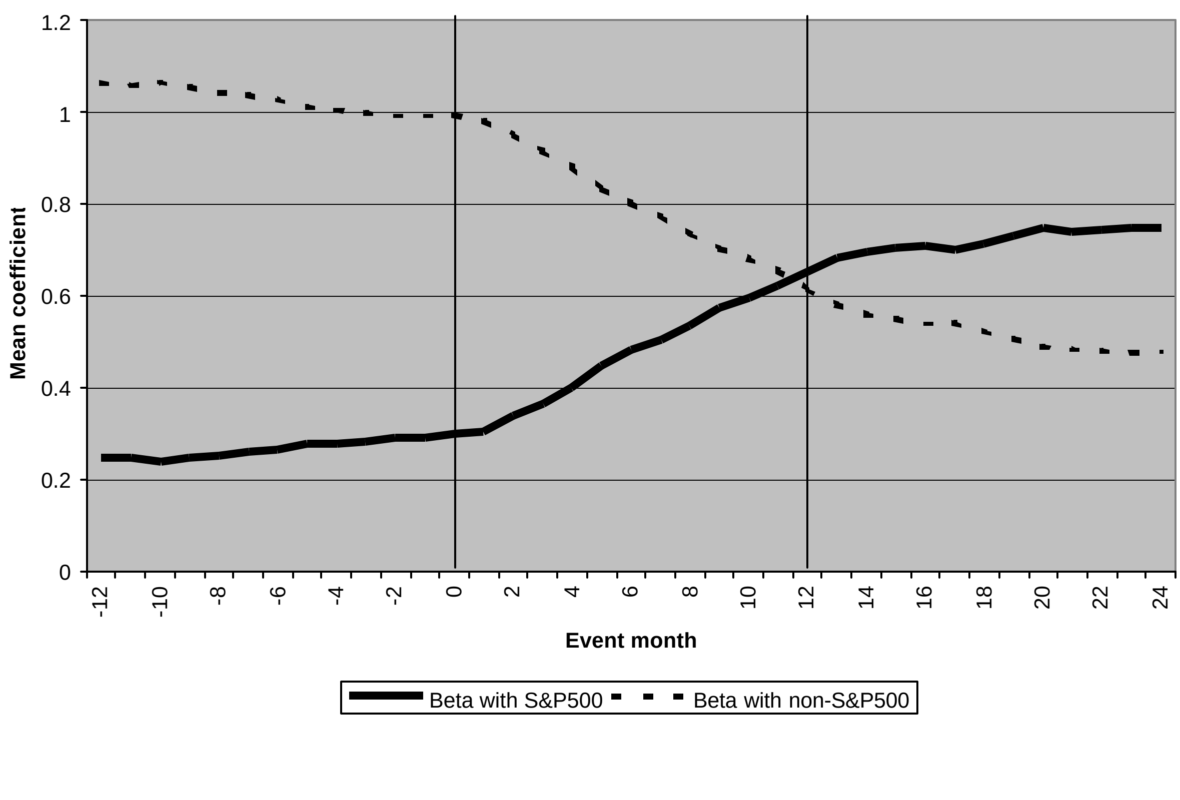
Panel B. Weekly Returns								
1995 – 1999 vs.								
	1970 – 1974	- 0.076	- 0.073	-0.050	- 0.032	- 0.026	- 0.132	0.002
	1975 – 1979	- 0.076	- 0.071	-0.050	- 0.030	- 0.025	- 0.110	0.002
	1980 – 1984	- 0.068	- 0.064	-0.044	- 0.020	- 0.016	- 0.105	0.002
	1985 – 1989	- 0.075	- 0.069	-0.050	- 0.033	- 0.029	- 0.054	0.090
	1990 – 1994	- 0.051	- 0.045	-0.029	- 0.012	- 0.008	- 0.041	0.136
Panel C. Monthly Returns								
1995 – 1999 vs.								
	1970 – 1974	- 0.115	- 0.103	-0.060	- 0.030	- 0.025	- 0.091	0.162
	1975 – 1979	- 0.114	- 0.102	-0.059	- 0.030	- 0.023	- 0.085	0.202
	1980 – 1984	- 0.102	- 0.090	-0.050	- 0.018	- 0.008	- 0.089	0.112
	1985 – 1989	- 0.116	- 0.105	-0.066	- 0.036	- 0.031	- 0.072	0.390
	1990 – 1994	- 0.095	- 0.083	- 0.045	- 0.017	- 0.012	- 0.061	0.274

Figure I. Changes in comovement of stocks added to the S&P 500 Index. Plots of the mean slope coefficients of bivariate regressions of returns of stocks added to the S&P 500 on returns of the S&P 500 Index and the non-S&P 500 rest of the market. The sample includes stocks added to the S&P 500 which were not involved in mergers or related events (described in the text), which have complete returns data over the entire event horizon examined in each figure (-12 to +24 months in daily and weekly returns data and –36 to +72 months in monthly returns data), and which remain in the Index for the full post-event horizon. For each added stock j, the bivariate model

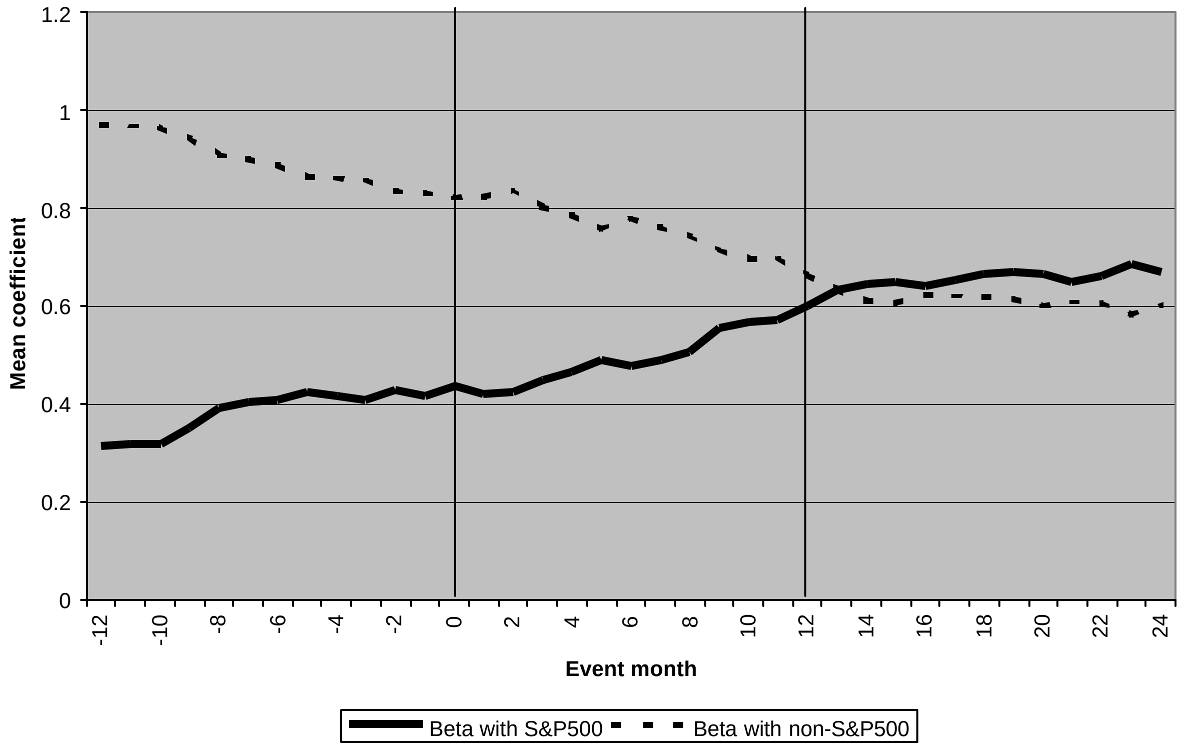
$$R_{j,t} = a_j + b_{j,SP500} R_{SP500,t} + b_{j,nonSP500} R_{nonSP500,t} + u_{j,t}$$

is estimated in rolling regressions where the sample intervals are [-12,-1] months for daily and weekly returns and [-36,-1] months for monthly returns. Returns on the S&P 500 (R_{SP500}) are from the CRSP Index on the S&P 500 Universe file. Returns on a capitalization-weighted index of the non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq are inferred from the identity described in Table 1. Returns from October 1987 are excluded. The mechanical influence of the added stock is removed, as appropriate, from both independent variables. The mean of each coefficient is plotted in event time. The left vertical line indicates the addition date; coefficients to the left of this line are estimated using only pre-event data. Coefficients to the right of the right vertical line are estimated using only post-event data. In between, coefficients are estimated using both pre- and post-event data. Panels A, B, and C show results for daily, weekly, and monthly returns, respectively.

A. Daily Returns (N = 384)



B. Weekly Returns (N = 384)



C. Monthly Returns (N = 211)

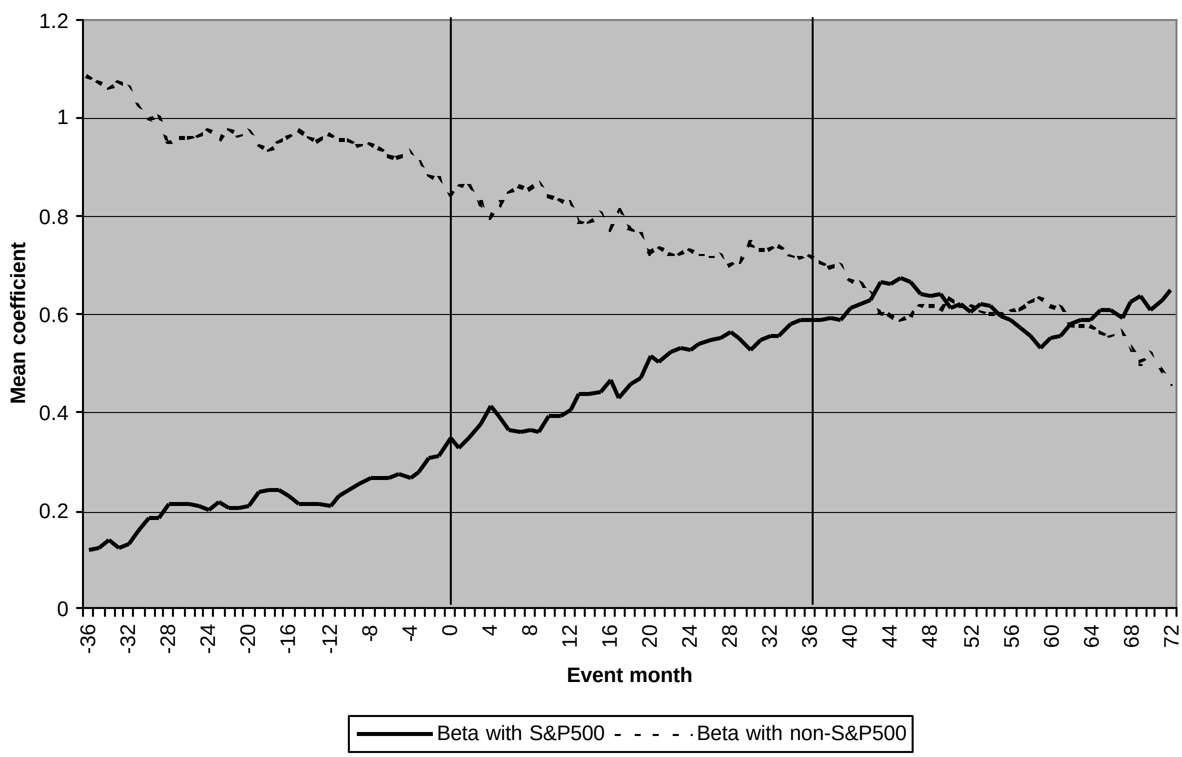
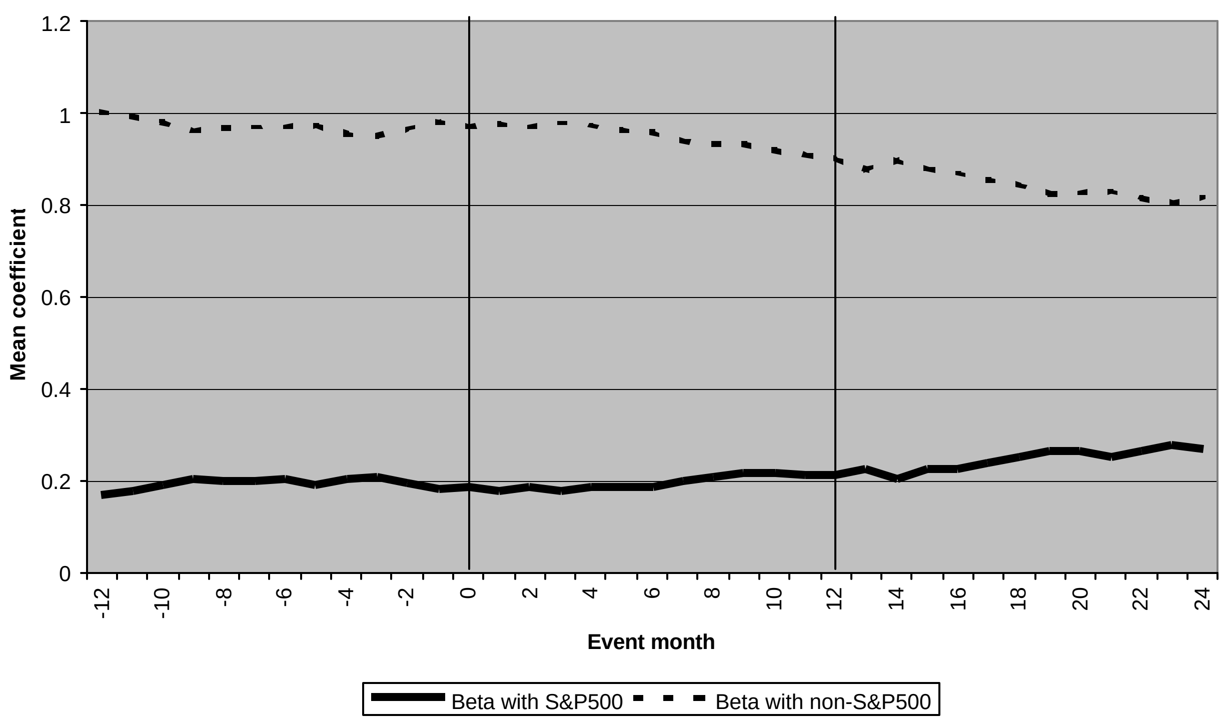


Figure II. Changes in comovement of stocks with matching characteristics to those added to the S&P 500 Index. Each stock in the event sample is paired with another stock which matches it on industry, market capitalization and growth in market capitalization over the pre-change estimation period (described in the text). The event sample includes stocks added to the S&P 500 which were not involved in mergers or related events, which have complete returns data over the entire event horizon examined in each figure (-12 to +24 months in daily and weekly returns data and –36 to +72 months in monthly returns data), which remain in the Index for the full post-event horizon, and for which a matching stock with complete data could be found. For each corresponding matching stock *j*, the bivariate model

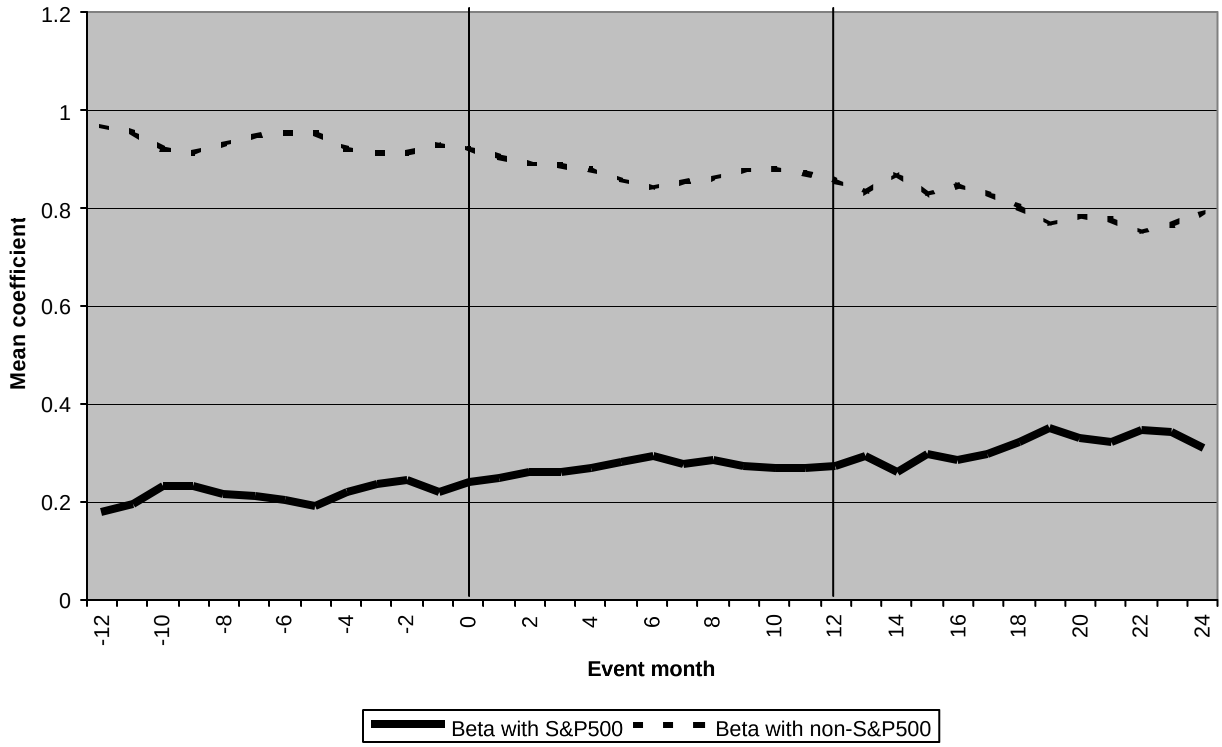
$$R_{j,t} = a_j + b_{j,SP500} R_{SP500,t} + b_{j,nonSP500} R_{nonSP500,t} + u_{j,t}$$

is estimated in rolling regressions where the sample intervals are [-12,-1] months for daily and weekly returns and [-36,-1] months for monthly returns. Returns on the S&P 500 (R_{SP500}) are from the CRSP Index on the S&P 500 Universe file. Returns on a capitalization-weighted index of the non-S&P 500 stocks ($R_{nonSP500}$) in the NYSE, AMEX, and Nasdaq are inferred from the identity described in Table 1. Returns from October 1987 are excluded. The mean matching stock coefficient is plotted in event time. The left vertical line indicates the addition date for the corresponding event stock; coefficients to the left of this line are estimated using only pre-event data. Coefficients to the right of the right vertical line are estimated using only post-event data. In between, coefficients are estimated using both pre- and post-event data. Panels A, B, and C show results for daily, weekly, and monthly returns, respectively.

A. Daily Returns (N = 322)



B. Weekly Returns (N = 322)



C. Monthly Returns (N = 144)