TRADE AND THE COMOVEMENT OF STOCK RETURNS: EVIDENCE FROM JAPAN

Robin M. Greenwood*
Harvard University
rgreenw@fas.harvard.edu

Nathan Sosner*
Harvard University
sosner@fas.harvard.edu

May 30, 2002

ABSTRACT

In April 2000, in one day, 30 stocks were replaced in the Nikkei 225 index in Japan. We analyze the change in comovement of returns of stocks added to and deleted from the index with the returns of stocks remaining in the index. A simple model shows that upon inclusion into (deletion from) a stock index, stocks should begin to comove more (less) with the index, due to a change in their trading pattern. The empirical findings provide sound support for these predictions: In the sample, daily index betas of the added stocks rose by an average of 0.60, while the average beta of the deleted stocks fell by 0.71. Our results confirm additional predictions of the model for changes in R^2, turnover, and the autocorrelation of returns upon index inclusion and deletion, and hold at daily, weekly and bi-weekly return horizons. Fundamentals based explanations fail to account for these findings. We conclude that correlated trading of index stocks causes excess comovement of stock returns. We argue that the distinct trading mechanism on the Tokyo Stock Exchange contributes to the significance and magnitude of our results.

* We are grateful to Nicholas Barberis, Malcolm Baker, Ken Froot, Emir Kamenica, Tom Knox, Mike Rashes, Jorge Rodriguez, Jeremy Stein, Jeff Wurgler, Tuomo Vuolteenaho, seminar participants at Harvard and especially John Campbell and Andrei Shleifer for helpful discussions. We remain responsible for any errors.

1. Introduction

Traditional finance models share the prediction that security prices move together only in response to correlated shocks to the expected cash flows or due to variation in discount rates. Indeed in a frictionless economy with perfectly rational investors, asset prices are merely a reflection of the underlying fundamental values. In such an economy, common variation in returns is fully accounted for by common variation in fundamentals and investor demand is irrelevant. In reality, however, we observe numerous examples of stock returns moving together either too little, or too much, to be justified by fundamentals, suggesting that the rational pricing postulate may be incomplete. For example, Pindyck and Rotemberg (1993) show that stock returns of firms in unrelated business lines comove significantly more than can be explained by common variation in the discount rates. Lee, Shleifer and Thaler (1991) show that discounts on U.S. closed-end funds tend to move together and moreover they comove with small stock returns. These studies and others[1] attribute excess comovement of stock returns to correlated demand shocks.

This paper presents evidence suggesting that at least at *short-horizons*, demand shocks which are not motivated by fundamentals are an important determinant of stock returns. We focus on a large redefinition of the Nikkei 225 stock index in Japan. In April 2000, the rules governing index membership were changed, causing the replacement of 30 stocks in the Nikkei index. As a result, returns on the stocks added to the index began to move substantially more with the returns on other index stocks while deleted stocks began to comove much less with the index stocks. To measure these changes in comovement, we estimate daily betas of the additions and deletions with respect to the equally weighted return of stocks remaining in the index after the redefinition. The betas of the additions rise by 0.60 while the betas of the deletions fall by 0.71. An alternative measure of comovement is the R^2 of the regressions of stock returns on the remainder index return. Indeed, the average R^2 of the regressions estimated with daily data rises from 0.04 to 0.15 for the additions and falls from 0.28 to 0.19 for the deletions. Similar results obtain at weekly and bi-weekly return horizons. In other words, the mispricing shows little tendency to revert as the return horizon increases. More importantly, the Nikkei additions and deletions are not small illiquid stocks – the additions represent more than 20% of the market capitalization of the Tokyo Stock Exchange, and the deletions, while smaller, were between fifth and eighth size deciles of the Tokyo Stock Exchange at the time of redefinition. In addition, changes in the comovement of turnover of additions and deletions with the turnover of the Nikkei 225 stocks are consistent with the view that trading is responsible for the short-run comovement of stock returns.

A common way to measure the effects of demand shocks has been to study the inclusion of stocks into stock indexes such as the S&P 500. Upon inclusion into a stock index, certain institutional investors are required to purchase the stocks. Shleifer (1986), Harris and Gurel (1986), Wurgler and Zhuravskaya (2002) document excess returns of stocks upon inclusion into the S&P 500 index in the United States. The change in price is hard to attribute to an information shock, since index inclusion is based on past performance and does not convey new information about the future dividends of the stock. Kaul, Mehrotra and Morck (2000) verify that returns around index redefinitions do not come from information by studying event returns associated with a mechanical change in the weights of the stocks already in the TSE 300 index in Canada.

The redefinition of stock indexes is also an appropriate setting to show that uninformed demand causes excess comovement of stock returns. Barberis, Shleifer and Wurgler (2002) (hereafter BSW) consider the comovement of stock returns in a model in which rational risk-averse arbitrageurs trade against investors who buy and sell securities based on some arbitrary rule unrelated to the underlying fundamentals. In the model, risky securities do not have perfect substitutes and arbitrageurs' ability to fully eliminate mispricing is limited by their risk aversion. As a result, arbitrageurs are unable to fully accommodate irrational demand without a change in the price. We follow this theoretical framework and model an event in which some index securities are replaced by previously non-index securities. The central assumption of our model is that there is a group of investors who trade only in index securities in proportion to index weights and irrespective of fundamentals. The model predicts that returns of securities added to the index should begin to comove more with the returns of securities remaining in the index. We measure comovement by the slope coefficient, or beta, in the regression of a security's return on the equally weighted return of securities remaining in the index, and the R^2 of this univariate regression. The model generates additional testable predictions for changes in the autocorrelations and cross-serial correlations of stock returns upon index inclusion and deletion, all of which apply directly to the Nikkei 225 redefinition and all of which are confirmed by our data. Taken together, the results are difficult to reconcile with *any* view in which fundamentals alone drive the short-run comovement of stock returns[2]. However, we are silent on the question of why some people trade the index rather than the full market portfolio. Index trading may be rational if there are differential trading costs or exogenous constraints on portfolio composition. Alternatively, performance of some institutional investors may be benchmarked to a stock index, in which case index membership of a stock, and not only its expected return and covariance with the other stocks may ultimately be part of an investor's utility function. Whatever the underlying cause of index trading, we show its prevalence and the economic significance of its effects.

Previous work on excess comovement of index stocks has focused on the effects of S&P 500 membership in the United States. In univariate regressions of stock returns on the S&P 500 index return, BSW find that the betas of additions rise by 0.12 while the betas of deletions fall by 0.15 during the sample period from 1976 to 1998. The increase in betas of additions is found to be much stronger in the second half of the sample after the introduction of the S&P 500 futures and options in 1982 and 1983, respectively. BSW interpret these results to be inconsistent with fundamentals driven comovement, but supportive of the view that excess comovement comes from correlated investor demand. Vijh studies the daily *market* betas of S&P 500 additions and finds that they increase by an average of 0.080 between 1975 and 1989. Similar to BSW, the Vijh results indicate that the increase in beta upon index inclusion becomes stronger following the introduction of the S&P 500 futures and options.

The evidence for the Nikkei 225 index provides important insights into the limits of arbitrage. First, it shows that models incorporating arbitrageurs with a finite risk bearing capacity and correlated demand shocks apply in markets with institutional arrangements different from the US stock exchanges, and thus are not merely suited to fit the empirical phenomena observed in the US.

Further, the changes in beta we measure are at least twice as large as those found in the U.S., and are significant enough to suggest that trading is responsible for *most* of the short-run comovement of stock returns in Japan. Theoretically, this may be due to a higher variance of demand shocks or to higher costs of arbitrage in Japan. Although we are unable to prove conclusively which of these is responsible for the magnitude of our results, we note that our findings are consistent with a low level of liquidity in Japan documented by a number of studies.[3] Unlike the NYSE, there is no designated market maker on the Tokyo Stock Exchange providing liquidity when there is an order imbalance. As a result, immediacy of transactions is sacrificed in favor of execution of trade at more favorable terms. Potential index arbitrageurs hence face the risk that their trades are not completed within a short time interval. In this setting, price effects of the demand shocks should be more pronounced. The fact that our results are much stronger and persist at longer horizons than those reported in United States is consistent with this hypothesis.

Finally, we consider the implications of our findings for estimation of CAPM beta. Since index demand appears to be a significant determinant of short-run stock returns, it raises the suspicion that OLS estimates of CAPM beta obtained with short-run return data may suffer from an omitted variable bias. Demand shocks for index stocks cause temporary mispricing of both the individual stocks and the market portfolio, so the OLS estimator of market beta obtained through regressing stock returns on a broad value-weighted market return is a biased estimator of the true loading on the market risk factor.[4] The bias may depend on the sampling of returns since the effects of demand shocks on the comovement of returns in turn depend on the horizon over which returns are measured. We explore the bias in OLS estimates of CAPM beta by studying multivariate regressions of stock returns on index returns and market returns at different horizons. Results of the multivariate regressions extend beyond the basic theoretical predictions considered in this paper as they demonstrate a changing pattern of comovement with respect to stocks outside of the index. The multivariate results allow estimation the bias in CAPM beta caused by the omission of demand related variable from the univariate CAPM regression. The bias that we measure is strong at both daily and weekly levels and, as expected, is especially strong for index stocks. For example, univariate CAPM beta estimates for the deletions before the event are upward biased by as much as 1.75. To summarize, although the April 2000 Nikkei event involves too few securities and too short return horizons for our results to be considered a rejection of the CAPM, they clearly suggest that demand shocks, and not the exposure to fundamental market risk, dominates short run variation in stock returns.

The next section outlines the predictions of a simple model of comovement. Section 3 describes the Nikkei 225 redefinition. Section 4 tests the predictions of the model on the comovement, autocorrelation and cross-serial correlation of stock returns and the comovement of turnover. Section 5 examines robustness issues. Section 6 describes how omitted demand shocks bias OLS estimates of CAPM beta. Section 7 concludes.

2. A model of comovement

This section presents a model that illustrates the effects of index trading on security returns. The model is developed within the theoretical framework put forward in BSW. With the introduction of some changes, this setup is directly applicable to the effects of the Nikkei 225 redefinition on stock returns and trading volume. Because our model does not break any new ground, we limit our discussion to the basic assumptions and testable hypotheses and leave the technical exposition for the appendix.

We assume a capital market that contains a risk-free security in perfectly elastic supply and $N+K$ risky securities in fixed supply paying uncertain liquidating dividends. N of these securities compose an equally-weighted security index, corresponding to the methodology employed by *Nihon Keizai Shimbun* in the construction of the Nikkei 225. Two types of agents operate in the market – index traders and arbitrageurs. Arbitrageurs are myopic investors with exponential utility of wealth. Index traders invest in the capital market in response to exogenous shocks to their wealth, but purchase only securities in the index fund in proportion to index weights. We do not address the rationality of the index traders' here. They may be rational traders with certain institutional constraints on their stock holdings or irrational traders who cannot solve the optimal portfolio choice problem, thus simply investing in the index. In this setup, all trading volume comes from interactions between index traders and arbitrageurs.

In each period a new piece of information about the liquidating dividend becomes available to investors. Information shocks are described by a two-factor model consisting of a market factor and an idiosyncratic factor. The covariance structure derived from the two factor model is referred to as the fundamental covariance. Importantly, there is *no* fundamental factor common to only index securities. In other words, there is no fundamental reason for index security returns to be more correlated than non-index security returns.

Capital market equilibrium is obtained through the market clearing of security demands of index traders and arbitrageurs. Since index trader demand is exogenous, price levels are determined by the willingness of arbitrageurs to absorb it. To obtain an analytical solution for return processes, we assume that the economy is in a covariance stationary equilibrium, in which arbitrageurs conjecture correctly the conditional covariance matrix of future stock returns. Under this assumption, returns are a linear function of the news about fundamentals and the index demand shock. The first proposition establishes a number of properties of the variances and covariance of security returns.

Proposition 1

(i) *The variance of an index security return is higher than the variance of a non-index security return and both are higher than the fundamental variance.*

(ii) *Covariance between any two index security returns is higher than the covariance between an index and a non-index security returns, which is in turn higher than the covariance between any two non-index security returns, and all three are higher than the fundamental covariance.*

The first part of Proposition 1 establishes that index trading induces excess volatility of security returns beyond fundamental volatility. Excess volatility of the index security returns is quite intuitive – index trading shocks are not fully diversified away by risk averse arbitrageurs and cause security prices to fluctuate more than can be justified by news about fundamentals. A more subtle result is that non-index securities exhibit excess volatility as well. This effect arises because arbitrageurs use *all* securities in the market to diversify idiosyncratic risks. Therefore, when an index trading shock occurs, the prices of the non-index securities are affected through the diversification strategies of the arbitrageurs, making their returns more volatile than can be explained by fundamentals. The second part shows that index trading shocks also induce excess covariance between any two securities.

Using the results of Proposition 1, we can show that although both index and non-index security returns respond to index demand shocks, these shocks have a stronger effect on the index security returns. This is because index trading shocks affect the returns of index securities directly.

To study changes in comovement of asset returns following redefinition of the Nikkei 225, we consider an event in which M index securities are replaced. We define the *remainder index return* as the equally-weighted return of all securities which were in the index before the redefinition and remained in the index after the event

$$\Delta P_{REM,t} = \frac{1}{N-M}\sum_{j=1}^{N-M}\Delta P_{j,t}$$

Proposition 2 states how the betas of securities that are added to and deleted from the index with respect to the remainder index should be affected by the redefinition event.

Proposition 2. *Consider a linear regression model in which return on a risky security j is regressed on the remainder index*

$$\Delta P_{j,t} = \alpha + \beta_{REM,j}\cdot\Delta P_{REM,t} + \xi_{j,t}$$

After redefinition

(i) The OLS estimate of $\beta_{REM,j}$ should increase for the securities added to the index and decrease for the securities deleted from the index.

(ii) The R^2 of this regression should increase for the additions and decrease for the deletions.

The change in beta is increasing in the coefficient of risk aversion and the variance of index shocks. These results are intuitive – arbitrageurs are less willing to bet against the mispricing when they are more risk averse and/or when the demand shocks are stronger. In addition to changes in contemporaneous correlations of returns, the model predicts that the autocorrelation and cross-correlation of returns should also be affected by index inclusion and deletion. Since fundamentals are not serially correlated, the only source of correlation is index trader demand. Demand is i.i.d over time, and so a high demand shock today implies that tomorrow's demand shock will on average be lower, and vice versa. This induces negative first-order serial and cross-serial correlation of returns.

Proposition 3.

(i) *Autocorrelations should become more negative for stocks added to the index and less negative for stocks deleted from the index.*

(ii) *Betas with respect to the leading and lagged remainder index should become more negative for the additions and less negative for the deletions.*

Of course, our simple model does not capture the whole richness of serial correlation patterns. For example, if cross-serial correlations are positive and sufficiently high, then the autocorrelation of portfolio returns may be positive while the autocorrelation of individual security returns is negative (see Lo and MacKinlay (1990)). In principle, one could generate complex time series patterns of this sort by introducing additional types of investors who react irrationally to past returns or cash flow news. One such model is described in Barberis and Shleifer (2002) in which investors switch between styles in response to past performance of the styles, generating serial correlation patterns that change with horizon. Their model predicts short-term positive autocorrelation, long term negative autocorelation, and negative cross-serial correlation. However, as we discuss below, in our data both individual security returns and index portfolio returns are weakly negative correlated.[5] Since these negative autocorrelations are consistent with the predictions of the model, we restrict the empirical analysis to this simple framework.

In addition to predictions related to return processes, the model yields simple implications for changes in the levels and correlations of trading volume of additions and deletions.

Proposition 4.

(i) *The volume of a security should increase upon addition and fall upon deletion.*

(ii) *Additions (deletions) to the index should experience increase (decrease) in the correlation of their volume with remainder index volume.*

We examine the empirical evidence for Propositions 1 to 4 in section 3.

3. The event: Nikkei 225 redefinition

The Nikkei 225 is the most widely followed stock index in Japan. The newspaper *Nihon Keizai Shimbun* (Nikkei) has maintained the index since 1970, following the discontinuation of the Tokyo Stock Exchange Adjusted Stock Price Average. The index includes 225 stocks, selected according to composition criteria set by Nikkei. Although the index guidelines are strict, changes to index composition prior to the redefinition were infrequent – typically one or two stocks per year.[6] Since the composition of the index had remained relatively fixed while the industrial composition of the stock market was changing, the Nikkei had become less correlated with the market over time, falling from a daily return correlation of 95% in 1998 to 84% in the first quarter of 2000. With the aim of reviving the relevance of the index, on April 14, 2000, Nikkei announced that changes in the "industrial and investment environments necessitated revision of rules covering selection of index components."[7] The change in criteria resulted in the substitution of 30 smaller issues with 30 large "New Economy" stocks. The revision became effective at the start of trading on April 24, 2000. The new index was in fact more representative of the market, as its correlation with the market increased to 95% following the event.

The one-time redefinition caused an enormous amount of trading during the week between the announcement and implementation. This trading can be approximately summarized as follows. The additions became a larger share (in Yen terms) of the new index than the deletions had taken in the old index. Accordingly, the weights of the stocks remaining in the index fell. The result was that investors tracking the index had to buy the additions and sell both the deletions and some fraction of the stocks remaining in the index. During the week after the announcement, many additions had returns exceeding 20% and many deletions fell by more than 30%. This is evident in Figure 1, where we plot returns surrounding the redefinition. Greenwood (2001) studies the cross-section of event returns in depth.

The value of the Nikkei 225 is determined by adding the ex-rights prices ($P_{i,t}$) divided by face value (FV_i) times a constant, dividing the total by the index divisor (D_t)

$$P_{Nikkei,t} = \frac{\sum_{i=1}^{225} \frac{P_{i,t}}{(FV_i / 50)}}{D_t}$$

Most stocks have a face value of 50, though some have face values of 5,000 or 50,000. The index divisor is adjusted daily to account for stock splits, capital changes, or stock buybacks[8]. Thus other than some dispersion in face value, the index is equally weighted. As a result, the index return is a price weighted average of index stock returns. Denoting returns of stock *j* in period *t* as $R_{j,t}$, the index return can be written as

$$R_{Nikkei,t} = \frac{P_{Nikkei,t} - P_{Nikkei,t-1}}{P_{Nikkei,t-1}} = \frac{D_t}{D_{t-1}} \sum_{j=1}^{225} \frac{\frac{P_{j,t-1}}{(FV_i / 50)}}{\sum_{i=1}^{225} \frac{P_{i,t-1}}{(FV_i / 50)}} R_{j,t}$$

Since $\frac{D_t}{D_{t-1}} \approx 1$, index returns are proportional to individual stock returns, with weights given by price over face value. In short, stocks with high prices disproportionately affect the index return. To prevent the results from being driven by the returns of stocks with high prices, in the subsequent analysis we study the equally weighted return. However, the results hold irrespective of which index we use.

Although we ignore index return weights for the rest of the analysis, it is useful to briefly describe the index weights and market values of the additions and deletions. First, the deletions are much smaller than the additions, the median market capitalization is Yen 25bn (about US $240mm), compared with Yen 978bn for the additions. Second, both additions and deletions had larger representation (i.e., index return weight) in the Nikkei 225 than they would have had in a market value weighted index.

Differences between additions and deletions are not limited to market value. Since the redefinition was intended to change the industrial composition of the index, the additions contain more electronics and financial firms while the deletions contain more chemicals and metal firms. Although the differences between additions and deletions are significant, we later verify that these differences do not drive the main result.

In our analysis, we focus on the returns and turnover of 28 additions and 30 deletions. We drop two additions, the Industrial Bank of Japan and Tokai bank, because of a subsequent takeover and delisting, respectively. Finally, when forming the remainder return index, we drop 9 of the 195 remaining stocks because of delisting, index removal, or because they were not in the index for long enough (more than one year) prior to the event.

4. Results

4.1. Changes in index beta after redefinition

As measures of comovement we use the index beta and R^2 of a univariate regression of log stock returns on the log remainder index return. Ideally, the fundamental characteristics of the index would not affected by the redefinition. The composition of the Nikkei 225 changed after the event thus causing a change in its fundamental properties, as well as a mechanical change in the beta of all stocks with respect to the Nikkei. We therefore can get closer to the ideal experiment by constructing an index that includes only the stocks that were in the Nikkei 225 before the event and remained in it after the event. We then analyze comovement by studying the beta and R^2 of additions and deletions with respect to an equally-weighted index of stocks remaining in the index, and not with respect to the Nikkei itself.[9]

Figure 2 shows the dramatic change in comovement of additions and deletions with the other stocks. Panel A shows the results for daily returns. We define a rolling window of 100 days and estimate regressions of log return of stock *i* on the remainder index log return

$$r_{it} = \alpha_i + \beta_{REM,i} r_{REM,t} + \varepsilon_{it} \tag{1}$$

For each 100-day window, we obtain separate beta estimates for each of the additions and deletions, then average separately across groups. We plot the time series of these group means.

High event returns of additions and deletions cause short-term trends in betas after the event. In Figure 2, the time series of betas exhibit sharp spikes during the event week[10]. However, the short-term trends in betas disappear once the rolling window leaves the event day. The dotted line in Panels A corresponds to the day on which the estimation window moves beyond the event day. From that day on mean betas remain fairly stable. Notably, the mean beta of the additions increases by than more than 0.50 and the mean beta of deletions drops by more than 0.60 when compared to average beta before the redefinition. Moreover, mean betas show no sign of reversion, on the contrary, the mean beta of additions continues to increase and the mean beta of deletions continues to decrease.

Panel B of Figure 2 repeats the above exercise with weekly returns and a 30-week estimation window. Beyond the fact that the change in betas is more pronounced for the deletions than for the additions, the results at the weekly level are essentially the same as at the daily level.

Are these changes in beta statistically significant? Table 1 shows basic tests of significance of the change in betas of additions and deletions. For each stock, we estimate two betas with respect to the remainder index – one for a window before the event and one for a window of equal length after the event. Then, as previously, we average betas across additions and deletions. In Panel A, the pre-event window includes returns from 100 trading days between November 26, 1999 and April 13, 2000. The post-event window includes returns from 100 trading days between May 1, 2000 and September 15, 2000. The results are again very strong. On average, the mean beta estimate of the additions goes up by 0.60, approximately a two-fold increase from its pre-event level. The mean beta of deletions goes down by 0.71. Both the increase in the betas of additions and the decrease in the betas of deletions are statistically significant at 1 percent confidence level.

The calculation of standard errors deserves special attention. The returns of additions and deletions exhibit a high degree of cross-sectional correlation. As a result, beta estimates are also cross-sectionally correlated. We use Seemingly Unrelated Regression approach to correct the t-statistic for this cross-sectional correlation in betas. Finally, to verify that a few outliers do not drive the observed changes in average betas, for each group we calculate the number and percentage of stocks whose beta goes up. We find that 86 percent of the additions (24 out of 28) experience increase in beta while 93 percent (28 out of 30) of the deletions experience declines in beta. We estimate the p-value on the hypothesis that betas were equally likely to go up as go down. For both additions and deletions, we reject this hypothesis at a 1 percent confidence level.

Panel B of Table 1 reports the change in beta measured using 250-day estimation windows before and after the event.[11] The results presented in Panel A survive the longer estimation window. The average beta of the additions increases by 0.46, while the average beta of the deletions decreases by 0.62. Both changes are significant at 1 percent confidence level. At this horizon, 71 percent (20 out of 28) of the betas of the additions go up, while 97 percent (29 out of 30) of the betas of the deletions go down.

To make sure that the changes in comovement are not merely the result of some trading frictions on a daily level, we increase the return horizon to weekly and bi-weekly.[12] Panel C shows that the change in average beta at the weekly level is commensurate with the changes at the daily level: betas of the additions go up by 0.55 while those of the deletions decline by 0.40. As shown in Panel D, the bi-weekly results are also very similar. This indicates that the change in comovement persists at least at the bi-weekly level and that there is little evidence of price correction in the short-run.

4.2. Changes in index R^2

Proposition 2 (ii) states that following inclusion into an index, the R^2 of the univariate regression of returns on the remainder index return should go up. Figure 3 plots the rolling R^2 from the regressions in (1), averaged separately across additions and deletions. The results are even more striking than those obtained with betas, as there is a sharp drop in the R^2 of deletions as soon as event returns leave the estimation window. The same is true for the additions, for which the R^2 increases by at least 10 percent when the event returns are no longer included in the estimation window. For both the additions and deletions, it appears that about half of their R^2 can be attributed to index membership. Panel B repeats the analysis using 30 weeks of weekly returns. The results appear stronger for the deletions but less pronounced for the additions.

Changes in R^2 reflect the economic importance of the index redefinition. The explanatory power of univariate regressions triples upon inclusion into the index and drops by half, following deletion from the index. We follow the same methodology as used in Table 1 and report average R^2 on the univariate regression specified in (1), using estimation windows of equal length before and after the event. In Panel A of Table 2, the average R^2 rises from 4 to 15 percent for the additions and falls from 28 to 19 percent for the deletions. Again, the calculation of standard errors deserves special attention, as returns and hence R^2 are likely to exhibit cross-sectional correlation. We use a parametric bootstrap[13] to compute standard errors, and find that for both additions and deletions, the change in R^2 is significant at a 1 percent confidence level. Motivated by the concern that the results are driven by a few outliers, we also report the fraction of stocks within each sample whose R^2 go up, together with the p-value on the null hypothesis that R^2 was equally likely to go up as go down. Panel A also shows that 86 percent (24 out of 28) of the R^2 statistics of the additions go up, while 80 percent (24 out of 30) of the R^2 statistics of the deletions go down.

Panel B of Table 2 reports the change in R^2 measured using 250-day estimation windows before and after the event.[14] The results presented in Panel A again survive the longer estimation window, and are in fact stronger. The average R^2 of the additions increases by 15 percent, while the average R^2 of the deletions decreases by 12 percent. Both changes are significant at 1 percent confidence level. We again check that these results are not driven by a few securities, and find that indeed 89 percent (25 out of 28) of the R^2 of the additions go up, while 97 percent (29 out of 30) of the R^2 of the deletions go down.

Panel C and Panel D report the change in R^2 measured using weekly and bi-weekly returns. The results remain strong for additions, while weakening for the deletions at the weekly level. Using bi-weekly returns, there is a 6 percent increase in R^2 for the additions and a 15 percent decrease for the deletions, with both changes significant at 10 percent confidence level.

So far, we have verified the two key predictions of the model, namely that the betas and R^2 of stocks included in (deleted from) a stock index rise (fall) with respect to the stocks in the index, and secondly that the R^2 of this regression goes up. The next part addresses the time series properties of returns.

4.3. The time-series properties of returns and non-synchronous trading

In the model, fundamental cash flows are independently distributed over time, so the only source of serial correlation of returns is the serial correlation of demand shocks. We showed earlier that i.i.d. index demand generates negatively autocorrelated demand shocks, and hence negatively autocorrelated returns. It follows that the autocorrelation of returns should become more negative for index inclusions and less negative for deletions. Table 3 explores this prediction by calculating the variance ratio statistics for daily and weekly returns. These statistics compare one period returns to returns measured over longer horizons. They are given by

$$VR_i(q) = \frac{\sum_{t=q}^{T}\left[\sum_{s=0}^{q-1}\left(r_{i,t-s}\right) - q\bar{r}_i\right]^2}{\sum_{t=1}^{T}\left[r_{i,t} - \bar{r}_i\right]^2}\left[\frac{T-1}{q(T-q-1)(1-q/T)}\right]$$

q denotes the aggregation value, T is the sample size, and $\bar{r}_i$ is the average return of stock i.[15] Values lower than one indicate negative autocorrelation of returns. Lo and MacKinlay (1999, p.54) provide an intuitive representation of the variance ratio

$$VR_i(q) - 1 \approx \frac{2(q-1)}{q}\hat{\rho}(1) + \frac{2(q-2)}{q}\hat{\rho}(2) + \ldots + \frac{2}{q}\hat{\rho}(q-1)$$

For example, $VR(2)$ in Table 3 is approximately one plus the first order autocorrelation of returns.

We estimate the variance ratio statistic before and after the event using 250 day windows. To eliminate the effect of the post-event abnormal returns on the autocorrelations, we increase the event window by 40 days.[16] The post-event period thus begins on June 26, 2000. The results are consistent with our basic predictions: the serial correlation of returns becomes more negative for additions and less negative for deletions. For the additions, the first-order autocorrelation decreases by 4 percent at the daily level and by 2 percent at the weekly level. While the former change is statistically significant, the latter is not. For the deletions, the first- order autocorrelation increases by 10 percent at the daily level and by 5 percent at the weekly level, both are highly statistically significant.[17] The larger increase in the variance ratios of the deletions for higher aggregation values indicates that higher order autocorrelations of the deletions have also become less negative. For the additions, at the weekly level the second order autocorrelation increases, but higher order autocorrelations become more negative, consistent with the view that they become exposed to negatively correlated index demand shocks. However, none of the weekly results are significant. At the daily level, higher order autocorrelations become more positive, but these changes are again small and insignificant.

In our discussion so far we interpreted the changes in autocorrelations within the framework of our model, but can these changes be explained within the efficient markets paradigm? In theory, there is a possibility that when securities are not continuously traded, their estimated autocorrelations are affected by a non-synchronous trading bias. In practice, however, we find it unlikely that the observed changes in the autocorrelations of the additions and deletions result from a time varying non-synchronous trading bias. First, Lo and MacKinlay (1999) demonstrate negative autocorrelations generated by non-synchronous trading are very small. For plausible changes in trading frequency, one can expect changes in autocorrelation substantially lower than 1 percent. Second, under the assumption that trading frequency should *increase* for additions and *decrease* for deletions, the autocorrelations should become *less* negative for additions and *more* negative for deletions, which is precisely the opposite of what we observe. While non-synchronous trading cannot explain the changes in the autocorrelations, index demand shocks can. Information shocks have permanent effects on security prices, while demand shocks are subsequently reverted by arbitrageurs. Therefore, as the magnitude of demand shocks relative to information shocks increases, securities will exhibit higher negative autocorrelation. This is exactly what appears in the data.

Although efficient market arguments such as non-synchronous trading cannot explain the changes in autocorrelation, can they explain some of the changes in beta? If returns used to estimate stock betas are measured at a high frequency, OLS beta may not be a consistent estimate of the true beta due to a non-synchronous trading bias. If the redefinition causes enough change in the frequency of trade, then changes in the *estimated* betas of additions and deletions may be merely the result of a non-synchronous trading bias. Scholes and Williams (1977) show that due to variation in trading frequency – or non-synchronous trading – the estimated betas of stocks traded very frequently or very infrequently will be biased downward, while the estimated betas of stocks with medium trading frequency will be biased upward. In this respect, if trading frequency of a stock is affected by index membership, then addition to or deletion from the index will affect the non-synchronous trading bias and, therefore, the estimated OLS beta. If index stocks are traded more frequently than non-index stocks, then OLS betas estimated with respect to the index should increase upon addition and decrease upon deletion. Since this scenario seems plausible, it becomes an empirical question how much it affects the main results.

Scholes and Williams derive an instrumental variables (IV) estimator which corrects for the non-synchronous trading bias in betas. They show that if non-synchronous trading effects are important at a one-day horizon, the consistent estimator of the true beta is given by

$$\hat{\beta}_{REM.,i} = \frac{b_{i,OLS}^{-1} + b_{i,OLS} + b_{i,OLS}^{+1}}{1 + 2\hat{\rho}_{REM}}$$

$b_{i,OLS}^{-1}$ is the OLS beta with respect to the lagged index return, $b_{i,OLS}$ is the conventional (and biased) OLS beta estimate, $b_{i,OLS}^{+1}$ is the OLS beta with respect to the lead index return, and $\hat{\rho}_{REM}$ is the estimate of the first order autocorrelation coefficient of the market return, in our case the remainder index return.

Corrected Scholes-Williams estimates of beta are given in Table 4. With daily returns, the decrease in beta estimates of deletions goes down by almost half. Consistent with the hypothesis that stocks become less frequently traded upon index deletion, the OLS beta estimates for deletions are almost identical to the Scholes-Williams IV estimates before the event, but suffer from a substantial downward bias after the event. Although the Scholes-Williams correction weakens the results, it does not come close to fully explaining the decrease in the betas of deletions. Not surprisingly, the non-synchronous trading bias seems to matter only when returns are measured at a daily frequency. With weekly data, there is very little change in the estimated betas after the correction. Both the pre- and post-event OLS betas of deletions estimated with weekly returns are somewhat downward biased, and there is little change in the bias after the event. There is no indication that a non-synchronous trading bias affects the additions. Using a 100-day estimation window, the increase in betas of the additions is not affected by the correction, while with a 250-day estimation window, the correction actually increases the measured change in beta. At the weekly level, the increase in the beta estimates of additions is somewhat smaller after the correction. However, since the non-synchronous trading effects are undetectable at the daily level, we doubt that non-synchronous trading is responsible for the result emerging at the weekly horizon.

Finally, the betas with respect to the lagged and the leading index returns require some interpretation. They indicate either a change in trading frequency or a change in the exposure to demand shocks. In the first case, lead and lag betas result from mismeasurement of returns due to non-trading and changes in these betas indicate changes in the frequency of trade. Alternatively, if index demand is i.i.d.

(as is assumed in the model) then both lead and lagged betas should decrease for additions and increase for deletions.[18] Table 4 suggests that index membership has a positive effect on trading frequency. For both 100-day and 250-day windows, betas of additions with respect to the lagged index, $b_{i,OLS}^{-1}$, substantially decrease, while their betas with respect to the lead index, $b_{i,OLS}^{+1}$, substantially increase. For deletions, in both the 100-day and 250-day windows, betas with respect to the lagged index increase substantially. For the 100-day return window, the beta of deletions with respect to the lead index decreases, while for 250-day window it remains essentially unchanged. Altogether, these observations admit possible non-synchronous trading effects on the betas estimated with daily returns. They also indicate that at the daily level the effect of non-synchronous trading on leading and lagged betas dominates the effect of the demand shocks. If the demand shocks were to drive these betas, then *both leading and lagged betas would change in the same direction*. For the additions, this occurs with weekly returns as both lagged and leading betas decrease. It is consistent with the view that additions become more subject to demand shocks since both their leading and lagged betas decrease. The evidence for deletions contradicts the index demand view since leading and lagged betas also decrease.

To summarize, the autocorrelations of returns change in the way predicted by the model. Non-synchronous trading does not appear to be an issue at the weekly return level and even at the daily return level, our main results survive.

One additional fact gives pause: while the turnover of the additions is virtually unchanged after the event, the turnover of the deletions goes up rather than going down. This implies that the downward bias in beta occurring after deletion from the Nikkei is because of *too frequent* trading rather than too infrequent trading. On the other hand, the strong increase of the beta of deletions with respect to the lagged index and decrease in their beta with respect to the lead index imply that the deletions are traded *too infrequently* after the event. In short, we should be careful with interpreting the Scholes-Williams results since they possibly introduce a considerable amount of noise into the beta estimates. More importantly, the results remain highly significant after the corrections.

4.4. *Turnover*

In the model, a change in the trading process of stocks results in a change in beta and R^2. The returns data strongly support this hypothesis. It is worthwhile to search for support for further evidence in the trading data. Ideally, we would observe the buys and sells of index traders directly. However, this is not feasible, and thus we study the implications of our model on trading volume. This section examines whether changes in the levels and comovement of turnover before and after the event are consistent with the model.

In the model, the turnover of non-index securities is always zero. Once added to the index, securities should experience an increase in their turnover. Of course, the model is not meant to capture the other reasons that market participants may trade.[19] For instance, the model cannot explain the fact that volume, as well as the standard deviation of volume, of the deletions went up after the event. This occurred as many index funds began to gradually unwind their positions in these stocks. Nonetheless, the prediction of the model that trading volume should become more correlated after index inclusion holds even if we add noise to the model causing trading volume in excess of that between arbitrageurs and index traders.

We follow Lo and Wang (2000) and use turnover (the number of shares traded divided by total number of shares) as a measure of volume – the number of shares traded divided by total number of shares. They show that turnover is well suited for studying the relations between volume and equilibrium asset pricing models. To measure changes in the correlation of turnover before and after the Nikkei redefinition, we use the same methodology that we used to examine the comovement of returns. We study the beta of turnover with respect to the average turnover of the stocks remaining in the index. Turnover is normalized by standard deviation to account for heteroskedasticity.

Table V presents estimates of turnover betas before and after the index redefinition. For each stock, we calculate the number of shares traded daily divided by the total number of shares normalized by standard deviation. The turnover index is defined as the equally weighted turnover of the remainders, henceforth *remainder turnover index*. We use the same estimation windows as before and measure "turnover betas" on the remainder turnover index prior to and after the event. For each stock *i*, we estimate the regression

$$Turn_{it} = \alpha_i + \beta_i \left[\frac{1}{N_{Rem}} \sum_{j \in Remainders} Turn_{jt} \right] + \varepsilon_{it} \tag{2}$$

Panel A presents the results using 100 days of turnover before and after the event. The average beta of additions rises from 0.74 to 1.15 while the average beta of deletions falls from 1.28 to 0.62. Since the independent variable is standardized, these betas can be interpreted as the sensitivity of stock turnover to a one standard deviation increase in the turnover of index stocks. The average R^2 of equation (2) rises by 10 percent for the additions and falls by 51 percent for the deletions. We correct all t-statistics for cross-sectional correlation as we did with the returns data.

The table shows similar results when the window is extended to 250 days or when we use weekly turnover. In short, the change in turnover following the event indicates that the changes in comovement of returns are similar to the changes in comovement of turnover.

We next ask whether the changes in turnover betas are driven by volume related to Nikkei 225 futures and options expirations. Futures and options on the Nikkei 225 index are traded on the Osaka Stock Exchange, and the expiration date is always the second Friday of the month. Although the futures contract is cash settled, the expiration causes significant increases in the volume of index stocks turnover due to liquidation of arbitrage portfolios combining stocks and derivatives. Table 6 shows the average daily turnover before and after the redefinition for additions and deletions. Average turnover on Mondays through Thursdays is virtually unchanged for the additions. However, their turnover increases dramatically on Friday. Since futures related trading on second Fridays is approximately double average daily trading during the rest of the month, we verify in Panel B of Table 5 that these outliers do not drive the changes in turnover betas. We estimate the same model as in Panel A, but exclude three days every month – one day before through one day after futures closes. This reduces the 100 day window by 13 days, for example. The results indicate that the changes in turnover betas do decrease substantially once we eliminate these days, falling from 0.40 to 0.23 for the additions and from –0.67 to –0.48 for the deletions. However, the results are still economically large and highly significant for both the changes in betas and in R^2. Similar results obtain when we use a 250-window.

In sum, the results for trading volume confirm our hypothesis that demand shocks significantly affect stock returns in the short run. Changes in comovement of the turnover are clearly indicative of changing patterns of trade. These patterns of trade are responsible for the strong increases in the betas of additions and decreases in the betas of deletions.

5. Robustness Tests

This section addresses several robustness issues. We first ask whether the returns associated with the characteristics of the additions and deletion drive the changes in index beta. We then ask whether non-random selection of deletions and additions introduces a selection bias into our results.

5.1. Do stock characteristics drive the results?[20]

A noted feature of the event was the difference in size and sector composition of the stocks comprising the additions and deletions. The additions were mainly large stocks that had experienced high positive returns during the past few years, and moreover were primarily in industries such as banking and electronics.[21] The deletions, on the other hand, were smaller stocks that had experienced low returns and declines in liquidity during the years prior to the event. They mostly represented stocks in chemicals, mining and metals industries. The differences in size and sector composition of additions and deletions are important only to the extent that the change in betas is driven by exposure to factors related to firm characteristics. The differences between the two groups observed in the data are not captured in the model, in which there is only one common risk factor affecting all security returns. However, if in reality there is variation in factor exposure between groups, it may affect the index beta estimates.

We begin with size. The median addition had a market capitalization of Yen 978 bn compared with Yen 25 bn for the median deletion (approximately $240mm at 105 Yen/$) and Yen 273 bn for the median stock in the remainders. Therefore, if there is a size factor in returns, the additions should carry much more exposure to it than the deletions. We obtain returns of Tokyo Stock Exchange small cap and large cap portfolios from Datastream. Taking the difference in log returns between these two indexes gives return on a portfolio which is long in small stocks and short in large stocks. We then estimate the beta of this portfolio return with respect to the remainder index return, and find that it remains virtually unchanged after the event. It is also possible to check that within-sample variation in size does not drive our results. We split the additions and deletions into those above and below their respective median size, and repeat the analysis from Table 2. Using 100 days of data, both subsamples of additions and deletions experience significant changes in index beta.

The second concern related to characteristics of additions and deletions is that their respective cash flows may be exposed to different industry shocks. The most common industries in the sample are electronics, banking, chemicals, and metals. Out of 28 additions, 9 are electronics firms, while 6 are banking firms. Of the 30 deletions, 7 are classified under chemicals while 9 firms fall under mining, iron and steel, or nonferrous metals. To make sure that industry characteristics are not driving our results, we first look at betas of the relevant Tokyo Stock Exchange industry indexes with respect to the remainder index return, using industry portfolio returns from Datastream. These indexes typically include more than 100 firms, and more than 200 in the case of the electronics index. In any case, if characteristics drive our results, then the correlations of the chemical index and the iron and steel index with the remainder returns should decline after the event, while the correlation of the electronics index and the banking index with the remainders should go up.[22] We estimate correlations on 250 days of data before and after the event and find that in three out of the four cases, they go in the wrong direction. For the electronics, however, the correlation goes up, consistent with an increase in industry risk driving some of the change in beta of the additions. To verify that the increasing correlation between the electronics industry and the remainder index does not drive our main result, we omit electronics firms from the additions and recalculate average betas. We find that indeed the change in beta declines but remains significant. Moreover, of the 19 securities left in the sample, 16 experience increases in beta.

5.2. Estimation windows

An additional set of robustness issues concerns the length and starting date of the estimation window before and after the event. First, as we noted earlier, an event return window of at least 10 days is excluded from all calculations since the returns surrounding the event are likely to overwhelm any measures of comovement we construct. However, Greenwood (2001) shows that in addition to large and significant event returns, the stocks in the index experienced reversion of event returns over a period up to 10 weeks after the event. Many arbitrageurs took large positions in the additions and deletions during the event. It is plausible to think that much of the reversion of prices during the weeks following the event may be driven by *price pressure of arbitrageurs* unwinding their positions in a correlated manner. In this case, we might observe excess covariance because of arbitrageurs, rather than in spite of them. In this case, the roles of the arbitrageurs and index traders are reversed, as the price pressure comes from the trading strategy of the arbitrageurs. This concern is easily addressed however, as we can extend the blackout window by 40 days. This only strengthens the results.

The last issue we address is whether the change in beta is due to the unwinding of index arbitrage positions during Nikkei 225 futures and options expiration dates, on the second Friday of every month. In short, is it the case that a few days of returns are driving our results? We eliminate a three-day window surrounding these days (the second Friday of every month) from our data, and estimate changes in average beta for 100 trading day windows. The betas of the additions rise by 0.61 while those of the deletions rise by –0.66, again, virtually unchanged.

5.2. Selection bias

One of the most problematic features of index redefinitions is the potential for selection bias in the composition of the additions and deletions. Clearly, in the Nikkei 225 event, replacement is hardly random: deletions were small stocks and represented only a minor fraction of index capitalization, while the additions were newer companies that had experienced high recent returns and growth in market capitalization. The difference in composition of these stocks is unimportant to the extent that selection criteria do not drive our main result. Hypothetically, the story might go as follows: suppose that the period of time we study was one in which the Nikkei index was falling. To increase the relevance of the index, the organizers drop the extreme losers and add winner stocks. This requires that the prices of deletions would fall even more than the index during this time, while the prices of stocks outside the index – including the additions – would rise. Following the redefinition, the abnormal returns cease and prices revert to the mean. This story could easily account for the changing index betas of the additions and deletions.[23]

However plausible this explanation might be, the evidence points the other way. During the 100 days prior to the event, the remainders outperformed the Tokyo Stock Exchange value weighted index (TOPIX). Moreover, the deletions outperformed not only the remainders but also the additions during these 100 days. Finally, we verify that the index betas we measure are not abnormal by extending the pre-event window back to June 1997, the index betas of the deletions remain high while those of the additions remain low.

Finally, there exists an (unlikely) possibility that stocks were deleted from the index not because of past returns but because of index betas. There are a variety of techniques to address this concern. The simplest one is to study changes in comovement for other high beta stocks. This is done as follows. For each stock in the index, betas are estimated with respect to *all* index stocks (i.e., both the remainders and deletions). Then a control group of the 30 highest beta stocks in our sample *that did not include deletions* is selected. In each case, the index was recomputed to include only N-1 stocks. These stocks have an average index beta of 1.40 before the event and 1.24 after the event, compared with 1.53 and 0.87 for the deletions. In other words, the index betas of comparable high beta stocks fall by only 0.16, less than a quarter of the 0.66 for the deletions. The interpretation of these results is that although additions and deletions are not randomly selected, selection bias does not drive the main result.

6. The bias in estimates of CAPM beta

Since demand is such an important determinant of short-run security prices, it is natural to ask how these omitted demand shocks affect the validity of the CAPM beta, obtained from univariate regressions of stock returns on a broad market index. The evidence thus far indicates that demand shocks cause temporary mispricing which is systematic across stocks. Therefore, both individual stocks and the market portfolio are subject to correlated mispricing. But if this is the case, then the OLS estimator of market beta obtained by regressing stock returns on a broad value-weighted market return is not a consistent estimator of the true loading on market risk factor. Since the additions become more exposed to index trading demand shocks, estimated betas should be more biased following index inclusion. Equally, since for the deletions, exposure to index trading demand shocks goes down, estimated betas should be less biased following index inclusion.

To better understand the effects of demand shocks on CAPM estimates, we study the results of multivariate regressions including both the remainder index return and the market return, proxied by the TOPIX index. The TOPIX is a value weighted index of the stocks traded on the First Section of the Tokyo Stock Exchange. In these regressions, the remainder index serves as a proxy for the demand shocks. Table 7 shows the results of this regression:

$$r_{it} = \alpha_i + \beta_{REM,i} r_{REM,t} + \beta_{MKT,i} r_{MKT,t} + \varepsilon_{it} \tag{3}$$

One caveat is in place here. We find that in the three months prior to the redefinition, the returns of the Nikkei 225 move in the opposite direction of the TOPIX. Such behavior of returns is not repeated anywhere else in the sample and is likely to dominate our results, especially in the 100 day window. To reduce the effect of these abnormally divergent returns during this period, we extend our pre-event estimation window to 500 days for daily returns and 100 weeks for weekly returns. The post-event sample is 250 days and 50 weeks, respectively.

Table 7 provides a number of observations. First, the coefficient of the additions' returns on the remainder index return is close to zero in the pre-event regressions. This holds closely to our theory – the remainder index return captures index demand and should be less important for stocks that are not in the index. We see that this changes dramatically after the event: the coefficient on remainder index return goes from an insignificant –0.04 to a significant 0.42. The dramatic change in this coefficient demonstrates the importance of the demand shocks for the returns of stocks in the index. The same pattern holds with the weekly data. We also find an increase in R^2 following addition into the index, consistent with the model's prediction that index demand becomes a more important determinant of returns following inclusion. Finally, we note that there is a decline in beta with respect to the market, from 0.96 to 0.67. This indicates that not only do additions begin to co-move more with index stocks, but also they begin to co-move less with non-index stocks. This result holds at the daily level only.

The deletions present the same broad patterns as the additions. At both the daily and weekly level, the beta of returns of the deletions with respect to index returns drops substantially (by 0.64) following the event. More interestingly, their exposure to the index is very high prior to the event, confirming the importance of demand shocks for the deletions. There is a dramatic decline in R^2 after the event, again supportive of our claim that demand shocks are significant determinants of stock returns. Finally, we note the increase in market beta following the event. At the daily level, beta rises by 1.09. This is similar to the result we obtained with additions: not only do deletions begin to co-move less with index stocks, they co-move more with stocks outside of the index.

The finding that the index redefinition changes the comovement properties of additions and deletions with respect to non-index stocks falls outside of our model. It is, however, explained by a model in which some groups of investors shift resources between stocks in different styles. Mullainathan provides motivation for this type of categorization. Barberis and Shleifer (2002) derive the properties of stock returns when stocks fall into two styles. Their model predicts excess negative correlation between stocks in different styles and excess positive correlation of returns of stocks within a style. If there are only two styles – index stocks and non-index stocks, then the Nikkei 225 redefinition should cause deletions to co-move more with non-index stocks and less with index stocks. Additions should begin to co-move more with index stocks and less with non-index stocks. For the deletions, at least, this seems to be true, and the additions show some evidence of a "style effect" at the daily return level. However, although we do believe index-stocks may form an investment style, we find it unlikely that non-index stocks together define a style. To better understand the change in market beta for the deletions, we investigate pre- and post-event correlations of deletion returns with a number of different "style" returns. We find (unreported) that the increase in market beta after the event is partly driven by an increase in their beta with respect to other small stocks. A plausible explanation of this finding is that the index redefinition caused the deletions to be reclassified from the index stock category to the small stock category.

Since omitted demand shocks are statistically and economically significant in our multivariate specifications, it follows that omitting this variable from individual return regressions will cause a bias in the estimate of the market beta. The OLS estimator of market beta can be decomposed as follows

$$\tilde{\beta}_{MKT,i} = \beta_{MKT,i} + \beta_{REM,i} \cdot \tilde{\beta}_{MKT,REM} \tag{4}$$

$\tilde{\beta}_{MKT,REM}$ is the beta coefficient of univariate regression of the remainder index return on the market return, given by

$$\tilde{\beta}_{MKT,REM} = \frac{Cov(r_{REM,t}, r_{MKT,t})}{Var(r_{MKT,t})}$$

The first term on the right hand-side of expression (4) is the true market beta parameter and the second term is the asymptotic bias.

We estimate the beta of the remainder index return with respect to the market return, $\tilde{\beta}_{MKT,REM}$, and find it to be positive before and after the event in both daily and weekly samples. Since we expect the true beta of the demand shock to be positive, the bias in the beta estimates obtained in univariate regression should be positive. Table 8 shows that before the event, there is essentially no bias in the market beta estimates for the additions. The bias does, however, increase after the event. For the deletions, there is a high positive bias in the univariate regression betas both before and after the event. The bias strongly decreases after the event, indicating the declining importance of demand shocks following index redefinition.

Another way to see the importance of the demand shocks for stock returns is to study how the adjusted R^2 changes when we add the remainder index return to the regression. Adjusted R^2 increases when the contribution of the new variable to the fit of the regression more than offsets the correction for the loss of an additional degree of freedom (see Greene (1997, p.255)). For additions there is little evidence that adding the remainder index return to the regression improves the fit of the regression. Neither is there evidence that the fit of the regression improves after the event. For deletions, however, adding the remainder index return to the regression dramatically improves the fit of the regression and especially so in the pre-event sample. The adjusted R^2 in the daily sample increases by 36 percent and in the weekly sample it increases by 49 percent. The increase in the adjusted R^2 before the event is much stronger than after the event, consistent with the hypothesis that demand shocks are an important factor in stock returns.

To summarize, estimating the loading on the market risk factor by running a regression of stock returns on the market return is likely to yield biased estimates of the true market beta. The results show that this bias is especially high for index stocks.

7. Conclusions

This paper describes a simple model of excess comovement of stock returns and tests its predictions with data from a large index redefinition in Japan. The model predicts that if arbitrageurs are limited in their risk bearing capacity, then correlated index demand should cause index stocks to comove more than implied by fundamentals. The comovement of stock returns and turnover with the index should increase upon addition to and decrease upon deletion from the index. We also predict that the autocorrelation and cross-serial correlation of stock returns should become more negative upon addition and less negative upon deletion.

Our data provide remarkably strong support for the theory. Following the event, additions begin to comove much more and the deletions begin to comove much less with index securities. The changes in comovement are reflected in the slope coefficients from the regression of individual stock returns on the equally weighted returns of stocks remaining in the index, and in the R^2 from these regressions. We verify that the results are not merely driven by a non-synchronous trading effect, firm characteristics or selection bias. Our findings also confirm additional predictions for changes in turnover, autocorrelation, and cross-serial correlation upon index addition and deletion.

We find that excess comovement of Nikkei 225 stock returns is stronger than excess comovement of S&P 500 stock returns in BSW and Vijh (1994). It is plausible that the lower liquidity of the Tokyo Stock Exchange documented in Hamao (1992), Hamao and Hasbrouck (1995) and Lehmann and Modest (1994) imposes additional costs on arbitrageurs, making them less effective at absorbing index demand shocks. Nevertheless, it is important to keep in mind that the TSE is a well functioning and liquid market.

An alternative explanation is that index demand shocks have substantially higher variance in Japan, but this seems unlikely given the low level of total indexation to the Nikkei 225 index.[24] The message of this paper is hence that demand shocks are important determinants of stock returns, irrespective of the microstructure of the market.

Finally, the analysis raises the question of whether the standard regression of a stock return on the broad market return yields a consistent estimate of the loading on the market risk factor. The multivariate regressions demonstrate that the OLS estimates of the market beta may be severely biased by the omission of an index demand variable. This paper focuses on index demand, but more generally, systematic demand shocks of different types may make conventional beta estimates unreliable.

Appendix 1: A Model of Comovement – Details

A. Setup

The capital market contains a risk-free security and $N+K$ risky securities paying uncertain liquidating dividends $D_{i,T}$ in period *T*. *N* of these securities are members of an index fund – an exogenously defined portfolio of risky securities. . Each risky security is in fixed supply *Q*. The risk-free asset is in perfectly elastic supply and its net return is normalized to zero. The information flow regarding dividend $D_{i,T}$ is given by

$$D_{i,t} = D_{i,0} + \sum_{s=1}^{t} \varepsilon_{i,s} \;, \quad \text{for all } i \tag{A1}$$

in which the information shocks $\varepsilon_{i,t}$ are identically and independently distributed over time and are normal with zero mean and covariance matrix Σ_D. The information shocks $\varepsilon_{i,t}$ are given by a two-factor model

$$\varepsilon_{i,t} = \psi_M f_{M,t+1} + \sqrt{1-\psi_M^2}\, f_{i,t+1}\,, \quad \text{for all } i. \tag{A2}$$

$f_{M,t+1}$ and $f_{i,t+1}$ are the market and the idiosyncratic factors, respectively. All factors are normally distributed with zero mean and variance of one and are i.i.d over time. The idiosyncratic factors are orthogonal to the market factor and are uncorrelated among securities. From (A2) the covariance structure of the information process is given by

$$\left(\Sigma_D\right)_{ij} \equiv \operatorname{cov}\left(\varepsilon_{i,t+1}, \varepsilon_{j,t+1}\right) = \begin{cases} 1, & \text{if } i = j \\ \psi_M^2, & \text{if } i \neq j \end{cases} \tag{A3}$$

Two types of agents operate in the capital market – index traders and arbitrageurs. Arbitrageurs are risk-averse myopic investors with exponential utility of wealth. In every period *t*, arbitrageurs determine their demand for risky securities by solving the following problem

$$\max_{N} E_t\left[-\exp\left(-\gamma W_{t+1}\right)\right] \tag{A4}$$

$$\text{s.t. } W_{t+1} = W_t + N_t'\left[P_{t+1} - P_t\right]$$

W_t, P_t, and N_t are arbitrageurs' wealth, the vector of security prices, and the fundamental demand vector at period *t*, respectively. Let P_t and N_t be $N + K$ vectors in which the first *N* elements are associated with the index securities and the remaining *K* elements are associated with the non-index securities. The solution to this problem yields the vector of fundamental demand as a function of the conditional expectation of the conditional covariance of prices and the coefficient of absolute risk aversion γ

$$N_t = \left[\gamma Var_t\left(P_{t+1}\right)\right]^{-1}\left[E_t\left(P_{t+1}\right) - P_t\right] \tag{A5}$$

Index traders invest in the capital market in response to exogenous shocks to their wealth, but purchase only securities in the index fund, in proportion to index weight. In every period, index traders have a random demand for index fund shares, u_t, i.i.d normally distributed with mean μ and variance σ_u^2. Index trading demand is given by $\frac{1}{N}\iota \cdot u_t$, in which ι is an indicator vector that takes on the value of one for index securities and zero for non-index securities.

B. Equilibrium with Index Traders

The market clearing condition is

$$N_t + \frac{1}{N}\iota \cdot u_t = Q \tag{A6}$$

Substituting in the demands of index traders and arbitrageurs and rearranging, we obtain

$$P_t = E_t\left(P_{t+1}\right) - \gamma Var_t\left(P_{t+1}\right)\left[Q - \frac{1}{N}\iota \cdot u_t\right] \tag{A7}$$

We assume that the capital market is in a covariance stationary equilibrium, in which $Var_t\left(P_{t+1}\right) = V$ for every *t*. Imposing this assumption and solving forward,

$$P_t = E_t(D_T) + \frac{1}{N}\gamma V \iota \cdot (u_t - \mu) - (T - t)\gamma V \left[Q - \frac{\mu}{N}\iota \right] \tag{A8}$$

Substituting (A1) into (A8) and taking the first difference yields an expression for security returns

$$\Delta P_t = \varepsilon_t + \frac{1}{N}\gamma V \iota \cdot \Delta u_t + \gamma V \left[Q - \frac{\mu}{N}\iota \right] \tag{A9}$$

in which $\Delta P_t \equiv P_t - P_{t-1}$.

The covariance matrix of security returns, V, can be characterized without solving explicitly for its elements. All securities have the same exposure to systematic risk as described by the two-factor model in (A2), and thus the only source of variation in the covariance matrix comes from index trading. As a result, covariance properties will be the same for all securities within the two groups, but different between the groups. That is V is given by

$$\underset{(N+K)\times(N+K)}{V} = \begin{bmatrix} A & C \\ C' & B \end{bmatrix}$$

$$\underset{N\times N}{A} = \sigma_1^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_1 \\ \cdots & \cdots & \cdots & \cdots \\ \rho_1 & \rho_1 & \cdots & 1 \end{bmatrix} \qquad \underset{K\times K}{B} = \sigma_0^2 \begin{bmatrix} 1 & \rho_0 & \cdots & \rho_0 \\ \rho_0 & 1 & \cdots & \rho_0 \\ \cdots & \cdots & \cdots & \cdots \\ \rho_0 & \rho_0 & \cdots & 1 \end{bmatrix} \qquad \underset{N\times K}{C} = \rho_{01}\sigma_0\sigma_1 \mathbf{1}_N \mathbf{1}'_K \tag{A10}$$

σ_1^2 and ρ_1 denote the variance and the correlation coefficient between two index securities, σ_0^2 and ρ_0 denote the variance and the correlation coefficient between two non-index securities, and ρ_{01} is the correlation coefficient between an index and a non-index securities. The vectors $\mathbf{1}_N$ and $\mathbf{1}_K$ are vectors of ones of length N and K, respectively.

Substituting (A10) into (A9) and dropping constants we obtain the following expressions for security returns

$$\begin{aligned} \Delta P_{i,t} &= \varepsilon_{i,t} + \phi_1 \cdot \Delta u_t \qquad \text{where } i \text{ is an index stock} \\ \Delta P_{k,t} &= \varepsilon_{k,t} + \phi_0 \cdot \Delta u_t \qquad \text{where } k \text{ is a non-index stock} \end{aligned} \tag{A11}$$

with coefficients of the index demand shocks given by

$$\phi_1 = \gamma\left[\frac{1}{N}\sigma_1^2 + \left(1-\frac{1}{N}\right)\sigma_1^2\rho_1\right], \quad \phi_0 = \gamma\rho_{01}\sigma_0\sigma_1 \tag{A12}$$

We repeat proposition 1 from the text.

Proposition 1. *Let the return processes be given by (A11) and (A12), then*

(i) *The variance of an index security return is higher than the variance of a non-index security return and both are higher than the fundamental variance of one. That is* $\sigma_1^2 > \sigma_0^2 > 1$.

(ii) *The covariance between any two index security returns is higher than the covariance between an index and a non-index security returns which is in turn higher than the covariance between any two non-index security returns, and all three are higher than the fundamental covariance. That is* $\sigma_1^2\rho_1 > \rho_{01}\sigma_0\sigma_1 > \sigma_0^2\rho_0 > \psi_M^2$.

Proof

(i) Consider the variance and covariance terms comprising the equilibrium conditional covariance matrix *V*. For any two index securities *i* and *j*

$$Var_t\left(\Delta P_{i,t+1}\right) \equiv \sigma_1^2 = 1 + \gamma^2\left[\frac{1}{N}\sigma_1^2 + \left(1-\frac{1}{N}\right)\sigma_1^2\rho_1\right]^2\sigma_u^2 \tag{A13}$$

$$Cov_t\left(\Delta P_{i,t+1}, \Delta P_{j,t+1}\right) \equiv \sigma_1^2\rho_1 = \psi_M^2 + \gamma^2\left[\frac{1}{N}\sigma_1^2 + \left(1-\frac{1}{N}\right)\sigma_1^2\rho_1\right]^2\sigma_u^2 \tag{A14}$$

For any two non-index securities *k* and *l*

$$Var_t\left(\Delta P_{k,t+1}\right) \equiv \sigma_0^2 = 1 + \gamma^2\left[\rho_{01}\sigma_0\sigma_1\right]^2\sigma_u^2 \tag{A15}$$

$$Cov_t\left(\Delta P_{k,t+1}, \Delta P_{l,t+1}\right) \equiv \sigma_0^2\rho_0 = \psi_M^2 + \gamma^2\left[\rho_{01}\sigma_0\sigma_1\right]^2\sigma_u^2 \tag{A16}$$

The covariance between an index security and a non-index security is

$$Cov_t\left(\Delta P_{i,t+1}, \Delta P_{k,t+1}\right) \equiv \rho_{01}\sigma_0\sigma_1 = \psi_M^2 + \gamma^2\left[\frac{1}{N}\sigma_1^2 + \left(1-\frac{1}{N}\right)\sigma_1^2\rho_1\right]\left[\rho_{01}\sigma_0\sigma_1\right]\sigma_u^2 \tag{A17}$$

(ii) Compare $Cov_t\left(\Delta P_{i,t+1}, \Delta P_{j,t+1}\right) \equiv \sigma_1^2 \rho_1$ and $Cov_t\left(\Delta P_{k,t+1}, \Delta P_{l,t+1}\right) \equiv \rho_{01}\sigma_0\sigma_1$

From (A14) and (A17)

$$\sigma_1^2 \rho_1 \left\{ 1 - \gamma^2 \sigma_u^2 \sigma_1^2 \left[\frac{1}{N\sqrt{\rho_1}} + \left(1 - \frac{1}{N}\right)\sqrt{\rho_1} \right]^2 \right\} = \rho_{01}\sigma_0\sigma_1 \left\{ 1 - \gamma^2 \sigma_u^2 \sigma_1^2 \left[\frac{1}{N} + \left(1 - \frac{1}{N}\right)\rho_1 \right] \right\}$$

For simplicity this can be rewritten as

$$\sigma_1^2 \rho_1 \{1 - a\} = \rho_{01}\sigma_0\sigma_1 \{1 - b\}$$

In the following steps we use the fact that $\rho_1 > 0$ which follows immediately from (A14). Notice that

$$a > b \Leftrightarrow \left[\frac{1}{N\sqrt{\rho_1}} + \left(1 - \frac{1}{N}\right)\sqrt{\rho_1} \right]^2 > \frac{1}{N} + \left(1 - \frac{1}{N}\right)\rho_1$$

Simplifying we get

$$a > b \Leftrightarrow \left[\frac{1 + (N-1)\rho_1}{N} \right]^2 \frac{1}{\rho_1} > \frac{1 + (N-1)\rho_1}{N}$$

From here

$$a > b \Leftrightarrow 1 + (N-1)\rho_1 > N\rho_1 \Leftrightarrow 1 > \rho_1$$

From (A13) and (A14) $\sigma_1^2(1 - \rho_1) = 1 - \psi_M^2$. By assumption $\psi_M^2 < 1$, ρ_1 is strictly less than 1. Therefore, *a* is strictly bigger than b. From (A14), $\sigma_1^2 \rho_1 > 0$ and $\sigma_1^2 \rho_1 \{1 - a\} = \psi_M^2 > 0$, therefore $\{1 - a\} > 0$. But then $a > b$ also implies $\{1 - b\} => 0$. From (A17), $\rho_{01}\sigma_0\sigma_1 \{1 - b\} = \psi_M^2 > 0$. Since $\{1 - b\} > 0$ $\rho_{01}\sigma_0\sigma_1 > 0$, and $\rho_{01} > 0$. Therefore

$$\sigma_1^2 \rho_1 > \rho_{01}\sigma_0\sigma_1 \Leftrightarrow \{1 - a\} < \{1 - b\} \Leftrightarrow a > b$$

The condition is satisfied and thus $\sigma_1^2 \rho_1 > \rho_{01}\sigma_0\sigma_1$.

(iii) Consider $Cov_t\left(\Delta P_{i,t+1}, \Delta P_{k,t+1}\right) \equiv \rho_{01}\sigma_0\sigma_1$ and $Cov_t\left(\Delta P_{k,t+1}, \Delta P_{l,t+1}\right) \equiv \sigma_0^2\rho_0$. In (ii) we have shown that $\rho_{01} > 0$, also (A16) immediately implies that $\rho_0 > 0$. From (A16) and (A17) and the fact that $\rho_{0,}\rho_{01} > 0$ it follows that

$$\rho_{01}\sigma_0\sigma_1 > \sigma_0^2\rho_0 \Leftrightarrow \frac{1}{N}\sigma_1^2 + \left(1 - \frac{1}{N}\right)\sigma_1^2\rho > \rho_{01}\sigma_0\sigma_1$$

And thus $\rho_{01}\sigma_0\sigma_1 > \sigma_0^2\rho_0$.

(iv) Now consider $Var_t\left(\Delta P_{i,t+1}\right) \equiv \sigma_1^2$ and $Var_t\left(\Delta P_{k,t+1}\right) \equiv \sigma_0^2$. From (A13) and (A15)

$$\sigma_1^2 > \sigma_0^2 \Leftrightarrow \left[\frac{1}{N}\sigma_1^2 + \left(1 - \frac{1}{N}\right)\sigma_1^2\rho_1\right]^2 > \left[\rho_{01}\sigma_0\sigma_1\right]^2$$

In part (ii) we show that $\rho_{1,}\rho_{01} > 0$, therefore

$$\sigma_1^2 > \sigma_0^2 \Leftrightarrow \frac{1}{N}\sigma_1^2 + \left(1 - \frac{1}{N}\right)\sigma_1^2\rho_1 > \rho_{01}\sigma_0\sigma$$

But then, by part (ii)

$$\frac{1}{N}\sigma_1^2 + \left(1 - \frac{1}{N}\right)\sigma_1^2\rho_1 > \sigma_1^2\rho_1 > \rho_{01}\sigma_0\sigma_1$$

Therefore $\sigma_1^2 > \sigma_0^2$.

(v) Finally from (A15) $\sigma_0^2 > 1$ and from (A16) $\sigma_0^2\rho_0 > \psi_M^2$. Using the results in parts (ii) – (iv), we obtain

$$\sigma_1^2 > \sigma_0^2 > 1 \text{ and } \sigma_1^2\rho_1 > \rho_{01}\sigma_0\sigma_1 > \sigma_0^2\rho_0 > \psi_M^2 \qquad \blacklozenge$$

Corollary 1. *The coefficient of index demand shock in the return equation is higher for an index security than for a non-index security, and both are positive. That is* $\phi_1 > \phi_0 > 0$.

Proof

Consider the expressions characterizing the demand shock coefficients in (A12). By part (ii) of Proposition 1

$$\phi_1 = \gamma\left[\frac{1}{N}\sigma_1^2 + \left(1-\frac{1}{N}\right)\sigma_1^2\rho_1\right] > \gamma\sigma_1^2\rho_1 > \gamma\rho_{01}\sigma_0\sigma_1 = \phi_0$$

and $\phi_1 > \gamma\psi_M^2 > 0$, $\phi_0 > \gamma\psi_M^2 > 0$. ♦

According to the assumptions, the information shocks $\varepsilon_{i,t}$ and $\varepsilon_{k,t}$ in return equations (A11) have a two-factor structure defined in (A2). Substituting this structure into the return equations we can define returns as a function of the market risk, the index demand shock and idiosyncratic shocks to dividends

$$\begin{aligned} \Delta P_{i,t} &= \psi_M f_{M,t} + \phi_1 \cdot \Delta u_t + \sqrt{1-\psi_M^2}\, f_{i,t} \quad \text{where } i \text{ an index stock} \\ \Delta P_{k,t} &= \psi_M f_{M,t} + \phi_0 \cdot \Delta u_t + \sqrt{1-\psi_M^2}\, f_{k,t} \quad \text{where } k \text{ a non-index stock} \end{aligned} \tag{A18}$$

C. Measuring Excess Comovement

This subsection analyzes changes in comovement of returns and turnover in a way directly applicable to the Nikkei 225 redefinition, by studying the index betas of security returns relative to the equally-weighted return of securities which remain in the index after redefinition.

Suppose that a redefinition of the index occurs in which M index securities are substituted with previously non-index securities. Define the remainder index return of the securities as the equally weighted return of all securities in the index which were in the index before redefinition and remained in the index after the redefinition

$$\Delta P_{REM,t} = \frac{1}{N-M}\sum_{j=1}^{N-M}\Delta P_{j,t} \tag{A19}$$

Proposition 2 establishes how betas of securities that were added and deleted from the index with respect to the remainder index should be affected by the redefinition event. It is restated in the text.

Proposition 2. *Consider a linear regression model in which the return on a risky security j is regressed on the remainder index*

$$\Delta P_{j,t} = \alpha + \beta_{REM,j} \cdot \Delta P_{REM,t} + \xi_{j,t} \tag{A20}$$

After redefinition

(i) *The OLS estimate of* $\beta_{REM,j}$ *should increase for the securities added to the index and decrease for the securities deleted from the index.*

(ii) *The* R^2 *of this regression should increase for the additions and decrease for the deletions.*

Proof

(i) Let $\Delta P_{i,t}$, $\Delta P_{k,t}$ and $\Delta P_{REM,t}$ denote index security, non-index security and remainder index returns, respectively. Consider the unconditional moments implied by expressions (A18) and (A19)

$$Cov\left(\Delta P_{i,t}, \Delta P_{REM,t}\right) = \psi_M^2 + 2\phi_1^2 \sigma_u^2 \tag{A21}$$

$$Cov\left(\Delta P_{k,t}, \Delta P_{REM,t}\right) = \psi_M^2 + 2\phi_1 \phi_0 \sigma_u^2 \tag{A22}$$

$$Var\left(\Delta P_{REM,t}\right) = \psi_M^2 + 2\phi_1^2 \sigma_u^2 + \frac{1-\psi_M^2}{N-M} \tag{A23}$$

OLS betas are then given by

$$\beta_{REM,i} = \frac{\psi_M^2 + 2\phi_1^2 \sigma_u^2}{\psi_M^2 + 2\phi_1^2 \sigma_u^2 + \frac{1-\psi_M^2}{N-M}} \text{ and } \beta_{REM,k} = \frac{\psi_M^2 + 2\phi_1 \phi_0 \sigma_u^2}{\psi_M^2 + 2\phi_1^2 \sigma_u^2 + \frac{1-\psi_M^2}{N-M}}$$

The change in beta of security *j* with respect to the remainder index upon index addition is thus given by

$$\Delta\beta_{REM,j} = 2\frac{\phi_1 \sigma_u^2 \left(\phi_1 - \phi_0\right)}{\psi_M^2 + 2\phi_1^2 \sigma_u^2 + \frac{1-\psi_M^2}{N-M}}$$

Symmetrically, the change in beta upon deletion is

$$\Delta\beta_{REM,j} = -2\frac{\phi_1\sigma_u^2(\phi_1 - \phi_0)}{\psi_M^2 + 2\phi_1^2\sigma_u^2 + \dfrac{1-\psi_M^2}{N-M}}$$

By Corollary 1, change in beta is positive for additions and negative for deletions.

(ii) Consider the unconditional variance of returns of an index and a non-index security

$$Var(\Delta P_{i,t}) = 1 + 2\phi_1^2\sigma_u^2 \tag{A24}$$

$$Var(\Delta P_{k,t}) = 1 + 2\phi_0^2\sigma_u^2 \tag{A25}$$

For univariate regressions, R^2 is the square of the correlation coefficient beween the dependent variable and the regressor. Let R_1^2 and R_0^2 denote the R^2 of the index return and the non-index return regressions, respectively. Then, by expressions (A13) – (A25)

$$R_1^2 = \frac{\left(\psi_M^2 + 2\phi_1^2\sigma_u^2\right)^2}{\left(\psi_M^2 + 2\phi_1^2\sigma_u^2 + \dfrac{1-\psi_M^2}{N-M}\right)\left(1 + 2\phi_1^2\sigma_u^2\right)}$$

and

$$R_0^2 = \frac{\left(\psi_M^2 + 2\phi_1\phi_0\sigma_u^2\right)^2}{\left(\psi_M^2 + 2\phi_1^2\sigma_u^2 + \dfrac{1-\psi_M^2}{N-M}\right)\left(1 + 2\phi_0^2\sigma_u^2\right)}$$

Define R^2 as a function of the demand shock coefficient in the return equations

$$R^2(\varphi) = \frac{\left(\psi_M^2 + 2\phi_1\varphi\sigma_u^2\right)^2}{\left(\psi_M^2 + 2\phi_1^2\sigma_u^2 + \dfrac{1-\psi_M^2}{N-M}\right)\left(1 + 2\varphi^2\sigma_u^2\right)}$$

Given that $\varphi \le \phi_1$, $\dfrac{dR^2(\varphi)}{d\varphi} > 0$

It follows that R^2 of this regression should increase upon addition to the index and decrease upon deletion from the index. ♦

It can be shown that the expression for the change in beta of a security added to the index with respect to the remainder index is given by

$$\Delta\beta_{i,REM} = \frac{2\frac{\phi_1}{\gamma}\left(\frac{\phi_1}{\gamma}-\frac{\phi_0}{\gamma}\right)}{\frac{1}{\gamma^2\sigma_u^2}\left[\psi_M^2\cdot\left(1-\frac{1}{N-M}\right)+1\cdot\left(\frac{1}{N-M}\right)\right]+2\left(\frac{\phi_1}{\gamma}\right)^2} \tag{A26}$$

Finally, the model yields simple implications for changes in the turnover, which we use as a measure of volume in the empirical analysis, of additions and deletions. The turnover of an index security is given by $\frac{1}{N}\cdot\frac{|\Delta u_t|}{Q_i}$ and the turnover of a non-index security is zero. The turnover of any two index securities is perfectly correlated, since in the model both are exposed to exactly the same demand shock. This result is restated in Proposition 4 in the text.

D. The time series properties of returns

The derivations presented in the subsection below are summarized in Proposition 3 in the text. News about fundamentals is not serially correlated, so the autocovariance functions of the security returns are given by

$$\begin{aligned}
\gamma_0(\varphi) &\equiv Var(\Delta P_{i,t}) = 1 + 2\varphi^2\sigma_u^2 \\
\gamma_1(\varphi) &\equiv Cov(\Delta P_{i,t}, \Delta P_{i,t-1}) = -\varphi^2\sigma_u^2 \\
\gamma_s(\varphi) &\equiv Cov(\Delta P_{i,t}, \Delta P_{i,t-s}) = 0, \quad \text{for all } s > 1
\end{aligned} \tag{A27}$$

$\varphi = \phi_1$ for index stocks and $\varphi = \phi_0$ for non-index stocks. By (A27) autocorrelations are given by

$$\rho_1(\varphi) \equiv \frac{\gamma_1(\varphi)}{\gamma_0(\varphi)} = \frac{-\varphi^2\sigma_u^2}{1+2\varphi^2\sigma_u^2}$$

$$\rho_s(\varphi) \equiv \frac{\gamma_s(\varphi)}{\gamma_0(\varphi)} = 0, \quad \text{for all } s > 1 \tag{A28}$$

Since from (A28) $d|\rho_1(\varphi)|/d\varphi > 0$ and by Corollary 1 $\phi_1 > \phi_0$, the first-order autocorrelation should become more negative upon addition to the index and less negative upon deletion from the index.

Next consider OLS betas with respect to the leading and lagged remainder index

$$b_{i,OLS}^{-1} = b_{i,OLS}^{+1} = \frac{-\phi_1^2\sigma_u^2}{Var(\Delta P_{REM,t})}, \quad \text{for all index securities}$$

$$b_{k,OLS}^{-1} = b_{k,OLS}^{+1} = \frac{-\phi_0^2\sigma_u^2}{Var(\Delta P_{REM,t})}, \quad \text{for all non-index securities}$$

Since $\phi_1 > \phi_0$, betas with respect to the leading and lagged index should become more negative for additions and less negative for deletions.

References

Barberis, N., and A. Shleifer, 2002, Style investing, *Journal of Financial Economics*, forthcoming.

Barberis, N., A. Shleifer, and J. Wurgler, 2002, Comovement, Mimeo University of Chicago.

Black, F., 1976. Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association*, Business and Economical Statistics Section, 177-81.

Bodurtha, J., D. Kim and C.M. Lee, 1995, Closed-end country funds and US market sentiment, *Review of Financial Studies* 8, 879-918.

Froot, K., and E. Dabora, 1999, How are stock prices affected by the location of trade?, *Journal of Financial Economics* 53, 189-216.

Greene, W., 1997, Econometric Analysis, 3rd Edition, Prentice-Hall, NJ.

Greenwood, R., 2001, Large events and limited arbitrage: Evidence from a Japanese stock index redefinition, Working Paper, Harvard University.

Hamao, Y., 1992, Tokyo Stock Exchange, The New Palgrave Dictionary of Money and Finance, 664-668.

Hamao, Y., and J. Hasbrouck, 1995, Securities trading in the absence of dealers: Trades, and quotes on the Tokyo Stock Exchange, *Review of Financial Studies* 8, 849-878.

Hardouvelis, G., R. La Porta, 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.

Lehman, B.N., and D.M. Modest, 1994, Trading and liquidity on the Tokyo Stock Exchange: A bird's eye view, *Journal of Finance* 49, 951-984.

Lo, A.W., and A.C. MacKinlay, 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3(1990), 175-206.

Lo, A.W., and A.C. MacKinlay, 1999, *A Non-Random Walk Down Wall Street*, Princeton University Press.

Lo, A.W., and J. Wang, 2000, Trading volume: definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257-300.

Kaul, A., V. Mehrotra, and R. Morck, 2000, Demand curves for stocks do slope down: New evidence from an index weights adjustment, *Journal of Finance* 55, 893-912.

Lee, C., A. Shleifer, and R. Thaler, 1991, Investor sentiment and the closed-end fund puzzle," *Journal of Finance* 46, 75-110.

Lewellen, John, 2002, Momentum and autocorrelation in stock returns, *Review of Financial Studies*, forthcoming.

Morck, R., B. Yeung and W. Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58(1).

Mullainathan, S., 2000, Thinking through categories, Working Paper, MIT.

Pindyck, R., and J. Rotemberg, 1993, The comovement of stock prices, *Quarterly Journal of Economics* 108(4), 1073-1104.

"Potential impact of the change in stock selection criteria for the Nikkei Average: High-priced tech stocks to be included", Nomura Securities Research Brief, April 21, 2000.

Roll, R., 1988, R^2, *Journal of Finance* 18(2), 541-566.

Shleifer, A., 1986, Do demand curves for stocks slope down? *Journal of Finance* 41, 579-590.

Vijh, A., 1994, S&P500 Trading strategies and stock betas, *Review of Financial Studies* 7(1), 215-251.

Wurgler, J., and K. Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks? *Journal of Business*, forthcoming.

Endnotes

---

[1] See also Froot and Dabora (1998), Hardouvelis, La Porta and Wizman (1994), Bodurtha, Kim and Lee (1995), Morck, Yeung and Yu (2000).

[2] Our results are incompatible even with semi-rational models such as Merton (1987) in which the discount rate depends on the investor clientele for the asset. Merton's model has no predictions for changes in R^2, for example. More importantly, it predicts increases in beta for stocks experiencing a reduction in their clientele, the opposite of what we find.

[3] Hamao (1992), Hamao and Hasbrouck (1995), and Lehman and Modest (1994) provide a detailed analysis of the trade on the Tokyo Stock Exchange. Hamao and Hasbrouck (1995), Lehman and Modest (1994) compare the trading mechanisms on the TSE and the NYSE

[4] Since the market portfolio includes both index and non-index stocks, the magnitude of the bias depends on the excess comovement with index and non-index securities. There is a possibility that demand shocks for index stocks cancel out simultaneous demand shocks for non-index stocks. This effect is related to "style investing" (Barberis and Shleifer (2002)) in which some groups of investors shift funds between stocks in different investment categories.

[5] Lewellen (2001) shows that negative autocorrelations and cross-serial correlations of monthly stock returns are responsible for the profitability of momentum strategies in the United States. In general, negative autocorrelations and cross-serial correlations are consistent with excess short-run comovement of stock returns. According to his argument, the source of profitability of momentum strategies comes from excess comovement rather than positive serial correlation of returns.

[6] Indeed we attempted to track down additional inclusions and deletions after 1990, but found that most of the deletions were subsequently delisted. Prior to 2000, there was an average of less than 2 stocks added or deleted each year. There are a few exceptions. In December 14, 1990 Nikkei announced a new Addition/Deletion standard effective October 1, 1991. This caused the substitution of 5 securities. Unfortunately, this event occurred during a time of especially volatile fundamentals- the collapse of the Japanese bubble between 1990 and 1992, during which time the Nikkei index fell by nearly 50%. Additional information on this redefinition and all changes in the composition of the Nikkei indexes can be found on the Nihon Keizai Net web page.

[7] A full description of index rules can be found at
http://www.nni.nikkei.co.jp/FR/SERV/nikkei_indexes/nifaq225.html#gen1

[8] The divisor was 10.18 on April 14, 2000.

[9] Since the actual index includes the deletions before the event and the additions after the event, our results on comovement are mechanically stronger with the actual index than with the remainder index.

[10] The explanation for these spikes is simple. Since the weights of the remainders decreased, institutional investors sold these stocks together with those of the deletions, driving the prices down. Simultaneously, they purchased shares of the additions, driving these prices up. During this short window, remainder returns were highly correlated with deletion returns and negatively correlated with addition returns. For a more complete explanation, see Greenwood (2002).

[11] Starting on May 1, 2000, we have a total of 284 days of returns. This limits the length of window we can study.

[12] Since the event happened in April 2000, we have only 18 months of returns following the event and hence limit our study to daily, weekly and bi-weekly horizons.

[13] We generate 2000 samples under the null hypothesis given by equation 1.

[14] Starting on May 1, 2000, we have a total of 284 days of returns. This limits the length of window we can study.

[15] For discussion of the statistical properties of this estimator see Lo and MacKinlay (1999, pp.47-83).

[16] The results are unchanged if we keep the original event window.

[17] The z-statistics in Table 2I are corrected for changing volatility (heteroskedasticity), but not for cross sectional correlation. Therefore, since the previous results suggest that security returns are positively correlated,

the z-statistics in Table 2I are potentially biased upward.

[18] Proof is in the appendix.

[19] Lo and Wang (2000) find at least 190 articles on volume. Within finance, they find that "volume is studied in several distinct subfields: market microstructure, price/volume relations, volume/volatility relations, models of asymmetric information and so on."

[20] The table for this section is available from the authors.

[21] See Appendix A2 for details on the sector composition of additions and deletions.

[22] There is no electronics index in the Tokyo Stock Exchange, however, the "electric machinery" index we use as a proxy for electronics is mainly composed of electronics firms. The Tokyo Stock Exchange web page contains detailed data on the current composition of firms in this industry.

[23] A related story can be told, and rejected, with leverage. With declining market value of equity, leverage mechanically increases *market* beta. See Black (1973). However, this does not necessarily imply changes in *index* beta, and moreover it has no prediction for R^2. In any case, the additions increased in price before the event and decreased after. The deletions increased after the event. If anything, correction for leverage would strengthen the results.

[24] Nomura Securities (2000) estimates the level of indexation to be about 1% of the total market value of Nikkei 225 stocks.

Figure 1. Cumulative Average Log Returns. Cumulative log returns of additions, deletions and remainders in an interval before and after the event, averaged separately across groups. Log returns are calculated on each day, and averaged across securities in each group. The figure shows cumulative returns based on a starting value of zero on April 14, 2000. The figure also shows cumulative log returns of the Tokyo Stock Exchange market value weighted index (Topix). The shaded area indicates a two-week period surrounding the event.

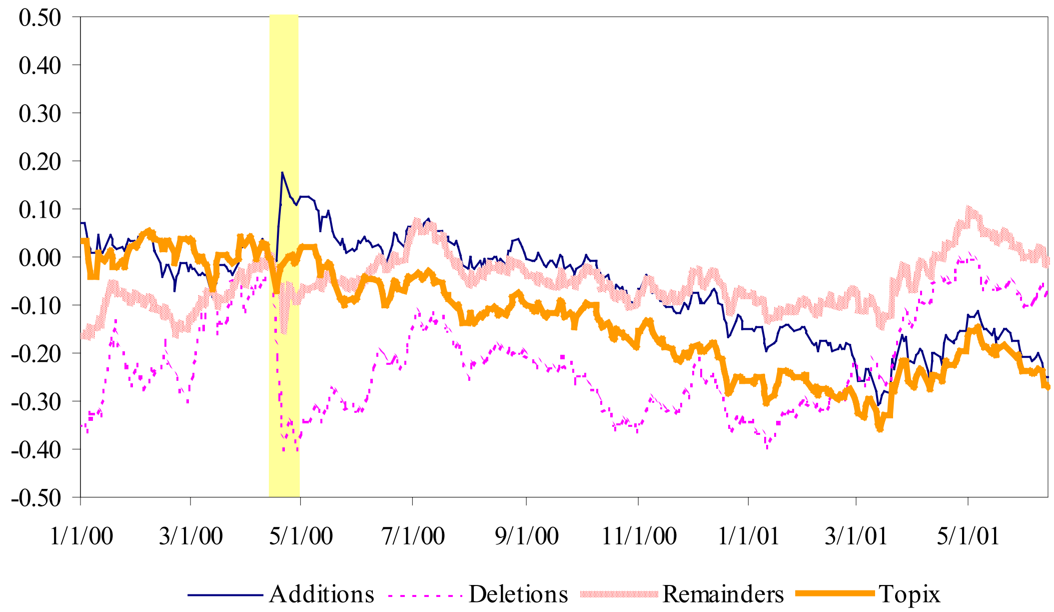
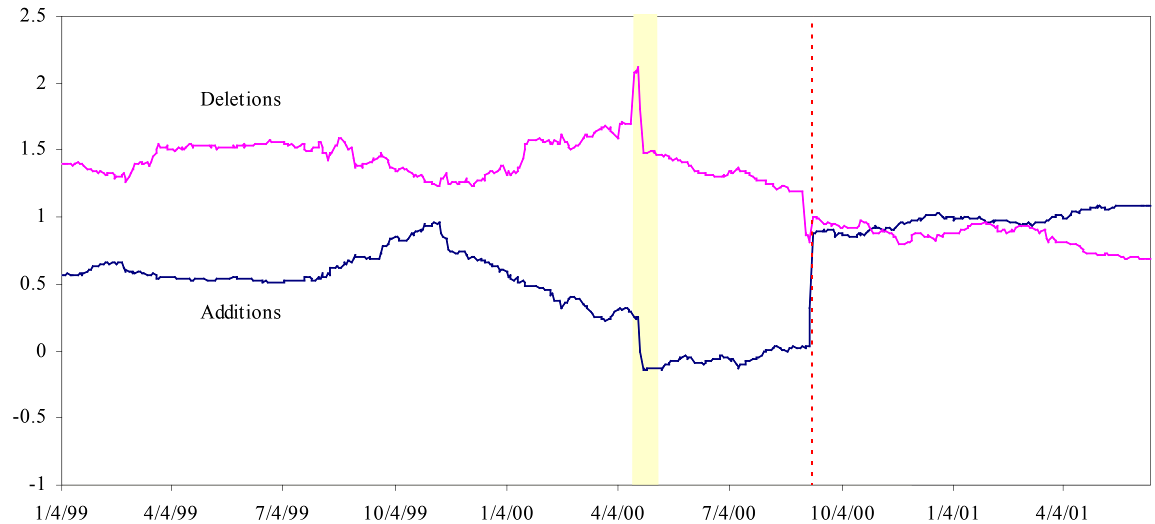


Figure 2. Rolling Betas. Average slope coefficients of rolling OLS regressions of a single stock return on a constant and the remainder portfolio return. The remainder index is formed as an equally-weighted portfolio of the stocks that remain in the Nikkei 225 index throughout the whole sample period. On each day, the slope coefficient estimates are averaged separately across additions and deletions. The time series of these cross sectional averages are then plotted. Panel A shows average rolling betas based on a 100 day window of daily returns. Panel B shows average rolling betas based on a 30-week window. In both panels, shaded areas denote a three-week period around the event – in this period, short-term changes in beta are dominated by event returns. A dotted line indicates the day (Panel A) or the week (Panel B) after which event returns are no longer included in the rolling estimation window.

Panel A. Daily Return Betas: 100 day rolling window



Panel B. Weekly Return Betas: 30 week rolling window

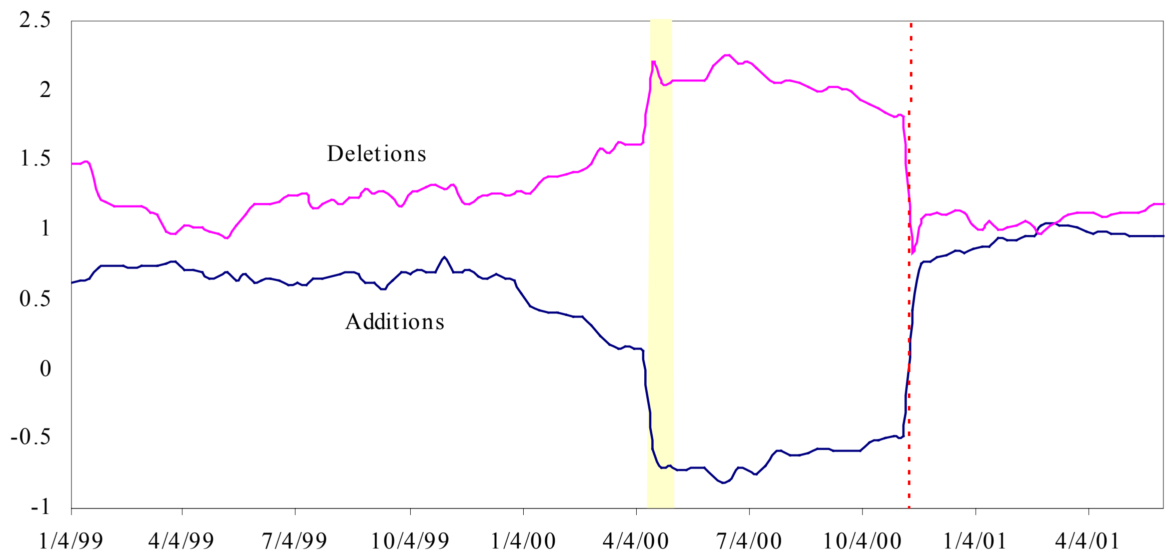
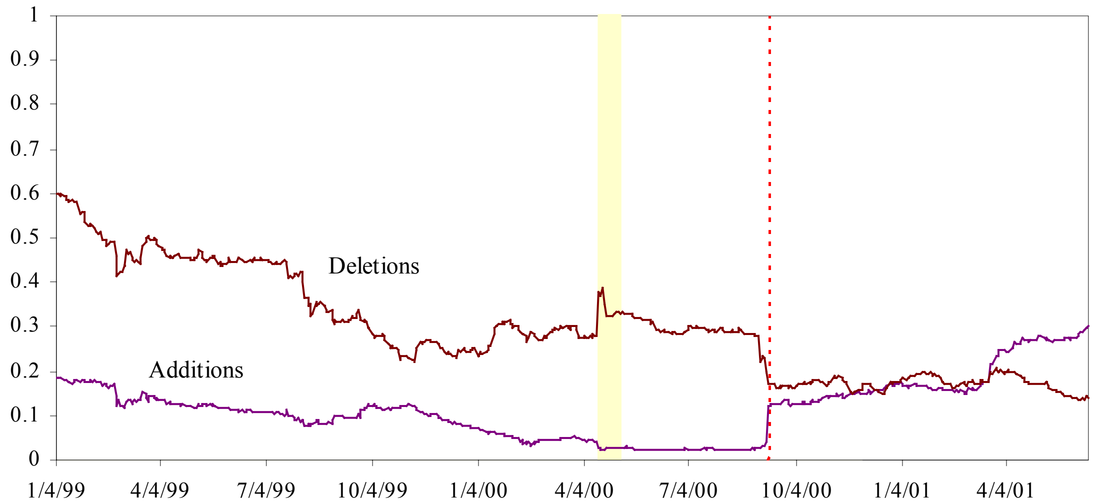


Figure 3. Rolling R^2. Average R^2 of rolling OLS regressions of a single stock return on a constant and the remainder portfolio return. The remainder index is formed as an equally weighted portfolio of the stocks remaining in the Nikkei 225 index throughout the whole sample period. On each day, the R^2 are averaged separately across additions and deletions. The time series of these cross sectional averages are then plotted. Panel A shows average R^2 based on a 100 day window of daily returns. Panel B shows average R^2 based on a 30-week window. In both panels, shaded areas denote a three-week period around the event – in this period, short-term changes in R^2 are dominated by event returns. A dotted line indicates the day (Panel A) or the week (Panel B) after which event returns are no longer included in the rolling estimation window.

Panel A. Rolling Daily R^2: 100 Day Rolling Window



Panel B. Rolling Weekly R^2: 30 Week Rolling Window

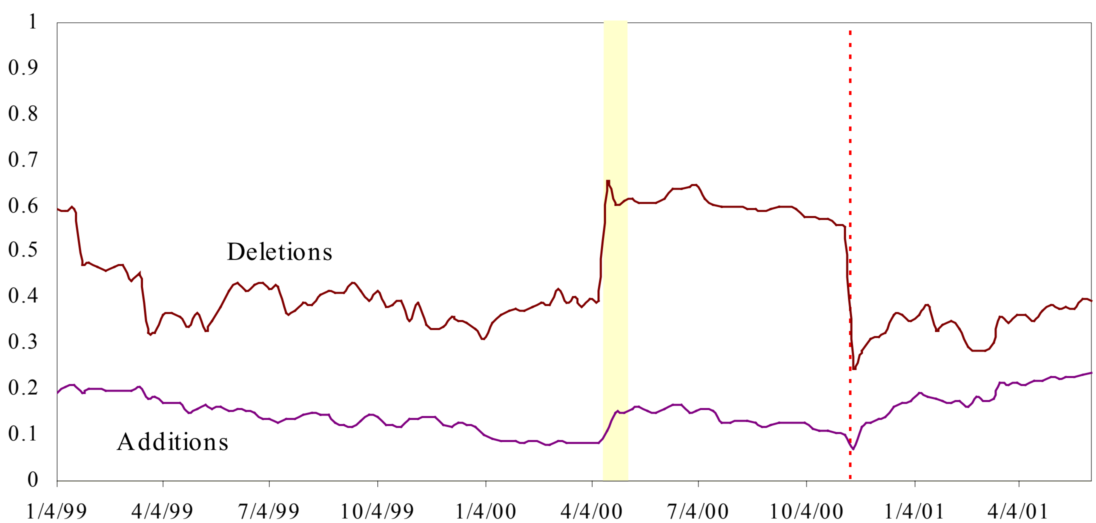


Table I. Changes in Index Beta of Stocks Added and Deleted from Nikkei 225. OLS regressions of log returns (r_{it}) on the equally-weighted log return ($r_{REM,t}$) of the stocks remaining in the Nikkei 225 index throughout the sample period.

$$r_{it} = \alpha_i + \beta_{REM,i} r_{REM,t} + \varepsilon_{it}$$

For each security *i*, we estimate the regression on a window of returns prior to the event, ending on April 14, 2000, the day of the redefinition announcement. We then estimate the regression on a window after the event, beginning on May 1, 2000. The table reports coefficients β_i before and after the event averaged across additions and deletions separately. The constant term is not reported. We omit two additions and 11 remainders due to delisting or merger during the pre- or post-event windows. We also calculate the number of firms within each group whose betas go up, and the p-value on the null hypothesis that betas were equally likely to go up as go down, reported in brackets. The first two panels estimate betas on daily data with window sizes of 100 and 250 days. The third panel uses weekly data for the estimation. The last panel uses 2-week log returns. T-statistics on the change in beta allow for cross-sectional correlation of the error term.

Sample	**N**	$\overline{\beta}_{Pre-Event}$	$\overline{\beta}_{Post-Event}$	$\Delta\overline{\beta}$	**[t-stat]**	**Fraction going up**	**[p-val]**
Panel A: Log Returns: 100 Day Window							
Additions	28	0.30	0.89	0.60	[3.97]	0.86	[0.00]
Deletions	30	1.70	0.99	-0.71	[-3.96]	0.07	[0.00]
Panel B: Daily Log Returns: 250 Day Window							
Additions	28	0.54	1.00	0.46	[6.10]	0.71	[0.01]
Deletions	30	1.48	0.85	-0.62	[-6.64]	0.03	[0.00]
Panel C: Weekly Log Returns: 50 Week Window							
Additions	28	0.33	0.89	0.55	[3.38]	0.68	[0.02]
Deletions	30	1.56	1.17	-0.40	[-1.89]	0.17	[0.00]
Panel D: Bi-Weekly Log Returns: 25 Observations Window							
Additions	28	0.27	0.75	0.48	[2.62]	0.54	[0.29]
Deletions	30	1.67	1.24	-0.43	[-1.88]	0.20	[0.00]

Table II. Changes in R^2 of Stocks Added and Deleted from Nikkei 225. OLS regressions of log returns (r_{it}) on the equally-weighted log return ($r_{REM,t}$) of the 184 stocks remaining in the Nikkei 225 index throughout the sample period.

$$r_{it} = \alpha_i + \beta_{REM,i} r_{REM,t} + \varepsilon_{it}$$

For each security i, we estimate the regression on a window of returns prior to the event, ending on April 14, 2000, the day of the redefinition announcement. We then estimate the regression on a window after the event, beginning on May 1, 2000. We report the average R^2 statistics before and after the event for additions to the index and deletions from the index separately. We also calculate the fraction of firms within each group whose R^2 go up, and the p-value on the null hypothesis that betas were equally likely to go up as go down. We omit two additions and 11 remainders due to delisting or merger during the pre- or post-event windows. The first two panels use daily data with window sizes of 100 and 250 days. The third panel uses weekly data and the last panel uses 2-week log returns. P-values on the change in R^2 allow for cross-sectional correlation.

Sample	**N**	$\overline{R^2}_{Pre-Event}$	$\overline{R^2}_{Post-Event}$	$\overline{\Delta R^2}$	**[p-val]**	**Fraction going up**	**[p-val]**
			Panel A: Daily Log Returns: 100 Day Window				
Additions	28	0.04	0.15	0.10	[0.00]	0.86	[0.00]
Deletions	30	0.28	0.19	-0.09	[0.00]	0.20	[0.00]
			Panel B: Daily Log Returns: 250 Day Window				
Additions	28	0.05	0.20	0.15	[0.00]	0.89	[0.00]
Deletions	30	0.28	0.16	-0.12	[0.00]	0.03	[0.00]
			Panel C: Weekly Log Returns: 50 Week Window				
Additions	28	0.07	0.17	0.10	[0.00]	0.75	[0.00]
Deletions	30	0.39	0.34	-0.05	[0.24]	0.36	[0.00]
			Panel D: Bi-Weekly Log Returns: 25 Observations Window				
Additions	28	0.07	0.14	0.06	[0.10]	0.57	[0.00]
Deletions	30	0.54	0.38	-0.15	[0.00]	0.18	[0.00]

Table III. Variance Ratio Statistics for Additions and Deletions. The table shows variance ratio statistics for daily and weekly log returns. VR(q) denotes the variance ratio computed for aggregation value q. The pre-event period is a 250 day period from September 20, 1999 to April 14, 2000; the post-event period is a 250 day period from June 26, 2000 to June 11, 2001. The z-statistics in parentheses are corrected for time-series volatility of returns.

	Pre-Event Additions	**Post-Event Additions**	**Difference**	**Pre-Event Deletions**	**Post-Event Deletions**	**Difference**
N	28	28	28	30	30	30
Panel A: Daily Variance Ratios						
VR(2)	1.02	0.98	-0.04	0.88	0.97	0.10
	[1.42]	[1.13]	[-3.62]	[9.18]	[1.81]	[9.71]
VR(4)	0.94	0.93	-0.01	0.82	0.93	0.11
	[2.03]	[2.73]	[-0.38]	[7.24]	[2.48]	[6.20]
VR(6)	0.88	0.90	0.03	0.78	0.90	0.13
	[3.15]	[2.93]	[0.98]	[6.90]	[2.78]	[5.35]
VR(8)	0.83	0.87	0.04	0.75	0.88	0.14
	[3.68]	[3.21]	[1.33]	[6.61]	[2.89]	[4.85]
VR(10)	0.81	0.85	0.04	0.70	0.89	0.19
	[3.73]	[3.31]	[1.20]	[6.78]	[2.30]	[6.02]
Panel B: Weekly Variance Ratios						
VR(2)	0.93	0.91	-0.02	0.91	0.96	0.05
	[2.53]	[3.38]	[-0.85]	[3.22]	[1.50]	[2.50]
VR(3)	0.89	0.88	-0.01	0.77	1.01	0.24
	[2.52]	[3.07]	[-0.48]	[5.78]	[0.32]	[8.77]
VR(4)	0.86	0.84	-0.02	0.71	1.05	0.33
	[2.66]	[3.26]	[-0.59]	[5.72]	[0.97]	[9.60]
VR(5)	0.85	0.81	-0.04	0.67	1.05	0.38
	[2.44]	[3.25]	[-0.96]	[5.64]	[0.96]	[9.45]

Table IV. Changes in Beta with Corrections for Non-synchronous Trading. Scholes-Williams (1977) regressions to correct for the presence of non-synchronicity in the recording of returns. For each security i , the table reports results of OLS regression of returns (r_{it}) before and after the event on index returns ($r_{REM,t}$), lagged index returns ($r_{REM,t-1}$), and leading index returns ($r_{REM,t+1}$):

$$r_{it} = \alpha_{-1i} + \beta_{-1,i} r_{REM,t-1} + \varepsilon_{it}$$

$$r_{it} = \alpha_{0i} + \beta_{0,i} r_{REM,t} + \varepsilon_{it}$$

$$r_{it} = \alpha_{+1i} + \beta_{+1,i} r_{REM,t+1} + \varepsilon_{it}$$

For each window, we estimate the auto-correlation of index returns, ρ . The corrected beta estimates are then given by

$$\hat{\beta}_{REM,i} = \frac{b_{i,OLS}^{-1} + b_{i,OLS} + b_{i,OLS}^{+1}}{1 + 2\hat{\rho}_{REM}}$$

The table reports coefficients for additions and deletions, averaged separately across groups. In the first panel, the stage 1 regressions are performed on 100 days of returns prior to and after the event. In the second panel, these regressions are repeated using 250 days of log returns; The third panel repeats the analysis using using 50 weeks of log weekly returns. In all panels, the last return day in the Pre-Event window is Friday April 14, 2000 and the first return day in the Post-Event window Monday, May 1, 2000. The last column reports the unadjusted differences in average beta taken from Table 1.

		Pre-Event Window					Post-Event Window					Differences	
Sample	**N**	$\hat{\rho}_{REM}$	$b_{-1,i}$	$b_{0,i}$	$b_{+1,i}$	$\hat{\beta}_i^{Corrected}$	$\hat{\rho}_{REM}$	$b_{-1,i}$	$b_{0,i}$	$b_{+1,i}$	$\hat{\beta}_i^{Corrected}$	$\Delta\hat{\beta}_i^{Corrected}$	$\Delta\hat{\beta}_i^{Uncorrected}$
Panel A: Daily Log Returns: 100 Day Window													
Additions	28	0.01	0.09	-0.05	0.30	-0.16	-0.07	0.66	-0.24	0.89	-0.08	0.58	0.60
Deletions	30	0.01	1.79	0.02	1.70	0.12	-0.07	1.35	0.19	0.96	0.00	-0.44	-0.71
Panel B: Daily Log Returns: 250 Day Window													
Additions	28	0.01	0.38	-0.03	0.54	-0.11	0.00	0.97	-0.06	1.00	0.03	0.59	0.46
Deletions	30	0.01	1.45	-0.04	1.48	0.03	0.00	1.13	0.23	0.85	0.04	-0.32	-0.63
Panel C: Weekly Log Returns: 50 Week Window													
Additions	28	0.06	0.25	0.04	0.35	-0.12	-0.11	0.60	-0.24	0.89	-0.19	0.35	0.56
Deletions	30	0.06	1.74	0.44	1.55	-0.04	-0.11	1.28	-0.01	1.13	-0.12	-0.46	-0.44

Table V*a*. Changes in Index Turnover Beta and R^2 (All Data). Average coefficients from OLS-regressions of daily and weekly turnover on the equally weighted turnover of stocks remaining in the Nikkei 225 index before and after redefinition.

$$Turn_{it} = \alpha_i + \beta_i \left[\frac{1}{N_{REM}} \sum_{j \in REM} Turn_{jt} \right] + \varepsilon_{it}$$

Turnover is defined as the volume in shares times the price of the stock divided by the market value, normalized by standard deviation. For each security *i*, we estimate the regression on a window of returns prior to the event, ending on April 14, 2000, the day of the redefinition announcement. We then estimate the regression on a window after the event, beginning on May 1, 2000. All available turnover data is used. Panel A reports estimates of beta, changes in beta, R^2 and changes in R^2 using turnover 100 days before and after the event. Panel B uses 250 days of data before and after the event. Panel C uses 50 weeks of weekly turnover data. T-statistics allow for cross-sectional correlation and are calculated using a parametric bootstrap.

Sample	**N**	$\overline{\beta}_{Pre}$	$\overline{\beta}_{Post}$	$\Delta\overline{\beta}$	**[t-stat]**	$\overline{R^2}_{Pre-Event}$	$\overline{R^2}_{Post-Event}$	$\overline{\Delta R^2}$	**[p-val]**
Panel A: Daily Turnover: 100 Day Window									
Additions	28	0.74	1.15	0.40	[5.37]	0.22	0.32	0.10	[0.45]
Deletions	30	1.28	0.62	-0.67	[-6.07]	0.62	0.11	-0.51	[0.00]
Panel B: Daily Turnover: 250 Day Window									
Additions	28	0.61	1.11	0.50	[11.25]	0.13	0.35	0.22	[0.00]
Deletions	30	1.24	0.54	-0.70	[-11.13]	0.52	0.09	-0.43	[0.00]
Panel C: Weekly Turnover: 50 Week Window									
Additions	28	0.58	0.98	0.39	[3.67]	0.11	0.27	0.15	[0.00]
Deletions	30	1.24	0.70	-0.54	[-3.32]	0.41	0.15	-0.27	[0.00]

Table V*b*. Changes in Index Turnover Beta and R^2 (Excluding Futures Closing Days). Average coefficients from OLS-regressions of daily and weekly turnover on the equally weighted turnover of stocks remaining in the Nikkei 225 index before and after redefinition.

$$Turn_{it} = \alpha_i + \beta_i \left[\frac{1}{N_{REM}} \sum_{j \in REM} Turn_{jt} \right] + \varepsilon_{it}$$

Turnover is defined as the volume in shares times the price of the stock divided by the market value, normalized by standard deviation. For each security *i*, we estimate the regression on a window of returns prior to the event, ending on April 14, 2000, the day of the redefinition announcement. We then estimate the regression on a window after the event, beginning on May 1, 2000. We omit three days of turnover surrounding the second Friday of every month, the closing day of Nikkei 225 options and futures contracts. Panel A reports estimates of beta, changes in beta, R^2 and changes in R^2 using turnover 100 days before and after the event. Panel B uses 250 days of data before and after the event. Panel C uses 50 weeks of weekly turnover data. T-statistics allow for cross-sectional correlation and are calculated using a parametric bootstrap.

Sample	**N**	$\overline{\beta}_{Pre}$	$\overline{\beta}_{Post}$	$\Delta\overline{\beta}$	**[t-stat]**	$\overline{R^2}_{Pre-Event}$	$\overline{R^2}_{Post-Event}$	$\overline{\Delta R^2}$	**[p-val]**
				Panel A: Daily Log Returns: 100 Day Window					
Additions	28	0.88	1.11	0.23	[2.71]	0.20	0.30	0.10	[0.49]
Deletions	30	1.11	0.63	-0.48	[-3.49]	0.30	0.12	-0.18	[0.03]
				Panel B: Daily Log Returns: 250 Day Window					
Additions	28	0.80	1.04	0.24	[4.40]	0.14	0.23	0.09	[0.01]
Deletions	30	1.01	0.70	-0.31	[-4.21]	0.24	0.11	-0.13	[0.00]

Table VI. Daily Turnover Before and After the Nikkei 225 Redefinition. For each security *i*, turnover is defined as the volume in shares times the price of the stock divided by the market value. We report average turnover for additions, deletions, and remainders, before and after the event. In the first two panels, we report the time series mean and standard deviation of group cross-sectional averages in parentheses. The last panel reports difference in average turnover before and after the event, and conventional t-statistics on the null hypothesis that the difference is zero.

Sample	**N**	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**2^{nd} Fridays**
				Pre-Event (%)			
Additions	28	0.20	0.21	0.22	0.22	0.25	0.28
		(0.08)	(0.08)	(0.08)	(0.07)	(0.09)	(0.12)
Deletions	30	0.22	0.23	0.24	0.27	0.46	1.07
		(0.14)	(0.11)	(0.13)	(0.12)	(0.50)	(0.72)
Remainders	184	0.19	0.20	0.21	0.22	0.28	0.45
		(0.09)	(0.07)	(0.08)	(0.07)	(0.17)	(0.25)
				Post-Event (%)			
Additions	28	0.21	0.22	0.23	0.23	0.31	0.54
		(0.08)	(0.06)	(0.07)	(0.09)	(0.18)	(0.19)
Deletions	30	0.41	0.42	0.43	0.39	0.41	0.52
		(0.29)	(0.29)	(0.26)	(0.22)	(0.26)	(0.32)
Remainders	184	0.20	0.22	0.24	0.23	0.28	0.45
		(0.08)	(0.07)	(0.07)	(0.08)	(0.15)	(0.16)
				Differences (Post-Event – Pre-Event) (%)			
Additions	28	0.01	0.01	0.01	0.01	0.06	0.26
		[0.36]	[0.45]	[0.52]	[0.68]	[1.94]	[8.63]
Deletions	30	0.19	0.18	0.19	0.13	-0.05	-0.55
		[4.35]	[3.33]	[3.94]	[3.29]	[-0.62]	[-4.94]
Remainders	184	0.02	0.02	0.02	0.00	0.00	0.00
		[2.37]	[3.40]	[3.14]	[0.51]	[-0.14]	[-0.27]

Table VII. Multivariate Regressions. OLS regressions of log returns (r_{it}) on the equally-weighted log return of the stocks remaining in the Nikkei 225 index and the market returns.

$$r_{it} = \alpha_i + \beta_{REM,i} r_{REM,t} + \beta_{MKT,i} r_{MKT,t} + \varepsilon_{it}$$

The market return is defined as the log return on the market-value weighted TOPIX index, which includes all stocks in Section 1 of Tokyo Stock Exchange. For each security *i*, the regression is performed on a window of returns prior to the event (ending April 14, 2000) and a window beginning 1 week after the redefinition became effective (starting May 1, 2000). To use as much data before the event as possible, *DoCoMo* is dropped from the additions, reflected in one fewer observations. The table reports average coefficients before and after the event for additions to the index and deletions from the index separately. The first panel estimates betas on daily data with a window size of 500 days before the event and 250 days after. The second panel uses 100 weeks of data before the event and 50 after. T-statistics are reported in brackets for the average change in beta and allow for cross-sectional correlation.

		Pre-Event			Post-Event			Differences		
Sample	**N**	$\overline{\beta}_{REM}$	$\overline{\beta}_{MKT}$	$\overline{R^2}_{Pre-Event}$	$\overline{\beta}_{REM}$	$\overline{\beta}_{MKT}$	$\overline{R^2}_{Post-Event}$	$\Delta\overline{\beta}_{REM}$	$\Delta\overline{\beta}_{MKT}$	$\Delta\overline{R^2}$
Panel A: Daily Log Returns										
Additions	27	-0.04	0.96	0.20	0.42	0.67	0.25	0.46	-0.28	0.04
		[0.62]	[15.88]		[6.99]	[11.20]		[7.60]	[-4.68]	
Deletions	30	2.05	-0.92	0.45	0.71	0.17	0.17	-1.34	1.09	-0.28
		[20.72]	[9.32]		[7.19]	[1.72]		[-13.53]	[11.04]	
Panel B: Weekly Log Returns										
Additions	27	0.03	0.87	0.24	0.21	0.78	0.27	0.18	-0.09	0.04
		[0.27]	[8.27]		[2.02]	[7.43]		[1.75]	[-0.84]	
Deletions	30	1.93	-0.84	0.54	1.28	-0.18	0.34	-0.64	0.66	-0.19
		[9.34]	[4.07]		[6.22]	[-0.88]		[-3.11]	[3.19]	

Table VIII. Decomposing the CAPM bias. The table reports average coefficients of market beta taken from univariate regressions of stock return on the market return, proxied by the log return on the Topix equally-weighted index. This is equal to the market beta from the multivariate regression of stock returns on the market return and index return, plus a bias term. The table also reports the Adjusted R^2 from each regression. The first two panels report estimates using 500 days of daily log returns prior to the event and 250 days of log returns after the event. We report average coefficients for additions and deletions separately, before and after the event. The bottom panels report estimates using 100 weeks of weekly log returns prior to the event and 50 weeks of log returns after the event.

	Univariate		Multivariate		Differences	
Period	**Market Beta in Univariate Regression** ($\tilde{\beta}_{MKT}$)	**Adjusted R^2**	**Market Beta in Multivariate Regression** (β_{MKT})	**Adjusted R^2**	**CAPM Bias**	**ΔAdj. R^2**
			Additions: Daily			
Before	0.93	0.16	0.96	0.19	**-0.03**	**0.03**
After	1.01	0.21	0.67	0.24	**0.34**	**0.03**
After - Before	**0.09**	**0.04**	**-0.28**	**0.04**	**0.37**	**0.00**
			Deletions: Daily			
Before	0.83	0.09	-0.92	0.45	**1.75**	**0.36**
After	0.74	0.12	0.17	0.16	**0.57**	**0.04**
After - Before	**-0.09**	**0.04**	**1.09**	**-0.29**	**-1.18**	**-0.33**
			Additions: Weekly			
Before	0.89	0.16	0.87	0.22	**0.02**	**0.06**
After	0.94	0.22	0.78	0.23	**0.16**	**0.01**
After - Before	**0.05**	**0.06**	**-0.09**	**0.02**	**0.14**	**-0.04**
			Deletions: Weekly			
Before	0.51	0.04	-0.84	0.53	**1.35**	**0.49**
After	0.76	0.16	-0.18	0.32	**0.94**	**0.16**
After - Before	**0.25**	**0.12**	**0.66**	**-0.21**	**-0.41**	**-0.33**

Table A1. Sector Composition of Additions and Deletions. Company name, sector, and Tokyo Stock Exchange identification code for each of the additions and deletions.

Code	Company	Sector	Code	Company	Sector
		Additions			
2914	Japan Tobacco	Foods	7270	Fuji Heavy Industries	Auto/Auto parts
4452	Kao	Chemicals	8035	Tokyo Electron	Trading Co
4505	Daiichi Pharmaceutical	Pharmaceuticals	8183	Seven-eleven Japan	Retailing
4523	Eisai	Pharmaceuticals	8264	Ito Yokado	Retailing
4543	Terumo	Pharmaceuticals	8267	Jusco	Retailing
6762	Tdk	Electronics	8302	IBJ*	Banking
6767	Mitsumi Electric	Electronics	8319	Daiwa Bank	Banking
6781	Matsushita Comm.	Electronics	8321	Tokai Bank+	Banking
6857	Advantest	Electronics	8355	Shizuoka Bank	Banking
6952	Casio Computer	Electronics	8403	Sumitomo Trust and Banking	Banking
6954	Fanuc	Electronics	8404	Yasuda Trust and Banking	Banking
6971	Kyocera	Electronics	8753	Sumitomo Marine and fire	Insurance
6976	Taiyo Yuden	Electronics	9020	East Japan Railway	Railroads/ Buses
6991	Matsushita Electric Works	Electronics	9433	DDI	Telecomm.
7211	Mitsubishi Motors	Auto/Auto parts	9437	NTT Docomo	Telecomm.
		Deletions			
1331	Nichiro	Fisheries	5331	Noritake	Glass
1501	Mitsui Mining	Mining	5351	Shinagawa Refractories	Glass
1503	Sumitomo Coal Mining	Mining	5479	Nippon Metal Indsutry	Iron and Steel
2108	Nippon Beet Sugar	Foods	5480	Nippon Yakin Kogyo	Iron and Steel
2601	Honen	Foods	5563	Nippon Denko	Iron and Steel
3104	Fuji Spinning	Textiles	5632	Mitubishi Steel Mfg.	Iron and Steel
3403	Toho Rayon	Textiles	5721	Shimura Kako	Nonferrous metals
4022	Rasa Industries	Chemicals	5805	Showa Electric Wire	Nonferrous metals
4064	Nippon Carbide	Chemicals	5981	Tokyo Rope Mfg.	Nonferrous metals
4092	Nippon Chemical	Chemicals	6461	Nippon Piston Ring	Machinery
4201	Nippon Synthetic Chemical	Chemicals	8061	Seika	Trading companies
4401	Asahi Denka Kogyo	Chemicals	8088	Iwatani International	Trading companies
4403	NOF	Chemicals	8236	Maruzen	Retailing
5105	Toyo Tire and Rubber	Rubber	9065	Sankyu	Land transportation
5302	Nippon Carbon	Chemicals	9302	Mitsui Soko	Warehousing

* Excluded from sample due to takeover in September 2000.
+ Excluded from sample due to delisting March 2001.