Working Paper #: 05-069

A CROSS SECTIONAL ANALYSIS OF THE EXCESS COMOVEMENT OF STOCK RETURNS*

ROBIN GREENWOOD
Harvard Business School

First Draft: September 2004
Revised: March 2005

* I thank Taro Hornmark, Nathan Sosner, and Tuomo Vuolteenaho for useful discussions.

Abstract

In the presence of limits to arbitrage, cross-sectional variation in periodic investor demand should be related to the degree of comovement of returns. I exploit the unusual weighting system of the Nikkei 225 index in Japan to identify cross-sectional variation in periodic demand for index stocks. Relative to their weights in a value weighted index, some stocks in the Nikkei are overweighted by a factor of ten or more. Using overweighting as an instrument for the proportionality between demand shocks for index stocks, I find a strong positive relation between overweighting and the comovement of a stock with other stocks in the index, and a negative relationship between index overweighting and comovement with stocks outside of the index. Put simply, overweighted stocks have high betas. The results suggest that excess comovement of stock returns is a consequence of an institutionalized commonality in trading behavior, rather than inefficiencies related to the speed at which index stocks incorporate economy-wide information.

Several recent papers show that security prices commove far in excess of their common fundamentals, casting doubt on the completeness of the rational pricing model in which comovement is fully explained by common variation in cash flows and discount rates. In their interpretation of the evidence, many authors have argued that excess comovement of stock returns may be explained by the price impact of correlated investor demand, or common *liquidity shocks* (e.g. Pindyck and Rotemberg, 1993; Lee, Shleifer and Thaler, 1991; Froot and Dabora, 1998).[1] Consistent with their intuition, there is growing evidence that comovement of security returns is empirically related to the trading patterns of certain groups of investors. Studying additions and deletions to the S&P 500 index, Barberis, Wurgler and Shleifer (2004) find that stocks tend to commove more (less) with index stocks after they are added to (deleted from) the index. Kumar and Lee (2003) show that correlated trades of retail investors are related to patterns of comovement in stock returns. However, while researchers have been able to identify the *existence of commonality* in investor demand, they have had less success determining whether the *degree of commonality* varies across securities, and if so, whether it bears any cross-sectional relation to comovement among security prices.

This paper develops cross-sectional predictions from a model in which the excess comovement of stock returns comes from correlated demand shocks, and tests them on a large panel of stocks between 1993 and 2003. The basic idea is as follows. Consider an economy in which there are 4 risky assets, A, B, C, and D in fixed and identical supply, and a risk-free asset in elastic supply. There are limits to arbitrage, so that uninformed demand for securities affects prices, at least in the short run. A set of investors regularly, and arbitrarily, buys and sells the risky assets in the fixed ratio of 1:1:2:-1, a weighting vector. Thus, when these investors buy one share of A, they also buy one share of B, two shares of C and sell one share of D. Following a positive investor demand shock, the prices of A and B rise, while the price of D falls. The price

---

[1] See also Vijh (1994), Hardouvelis, La Porta and Wizman (1994), De Long, Shleifer, Summers and Waldmann (1990), and Morck, Yeung, Yu (2000). Similar arguments have been made in the context of 'contagion' between stock markets (e.g. Calvo, 1999).

of C also rises, but more than the prices of A and B. Conversely, following a negative demand shock, the price of C falls more than the prices of A and B. After controlling for common fundamentals, the covariance of returns of A and C should exceed the covariance of returns of A and B. Similarly, the covariance between the returns of C and D should exceed the negative covariance between the returns of A and D. The key insight is that although the *level* of uninformed demand in any period is random, the proportionality between demand shocks generates a cross sectional relation between the weighting vector and the comovement of stock returns.

While easy to envision in theory, cross-sectional variation in regular demand shocks of the type described above can be difficult to find in practice. I use the unusual weighting system of the Nikkei 225 index in Japan to identify variation in regular demand shocks for index stocks, and relate this variation to patterns in comovement. The Nikkei 225 index is effectively equal weighted, meaning that stocks exert influence on the index return in proportion to their price. The index weights of some stocks exceed by an order of magnitude their weights in a value weighted index. Thus, when Nikkei 225 index investor demand rises (perhaps because of sentiment for Nikkei stocks, or perhaps because of market-wide sentiment), investors purchase significantly more of some index stocks than they would if they were using the value weighted TOPIX index as the benchmark instead. Conversely, when Nikkei 225 index investor demand falls, investors sell more of these stocks than they would if the index were value weighted. As of September 1, 2003, the Nikkei 225 weights of more than 50 stocks were less than their weight in the value weighted TOPIX, while the weights of 48 stocks were more than 5 times their TOPIX weight. If excess comovement of index stock returns is the result of uninformed demand for index stocks, then the returns of overweighted stocks should commove more with the equal weighted return of the other index stocks, while the returns of underweighted stocks should commove less.

I analyze the comovement between 298 Nikkei index stocks and 1458 non-index stocks between 1993 and 2003. In repeated cross-sectional regressions, I study the relation between comovement of a stock with other index stocks, and a measure of its overweighting in the Nikkei

225 index relative to the value weighted TOPIX. Index overweighting acts as an instrument for the unobserved true cross-sectional structure of demand.

The results provide strong support for my cross-sectional predictions. In the first set of tests, I detect a positive cross sectional relation in 25 out of 26 cross sections. Put simply, stocks that are overweighted in the index have much higher betas. I also study the comovement of index stock returns with the returns of stocks outside of the Nikkei 225 index, with converse results: index overweighting is significantly negatively related to the comovement of an index stock with stocks outside of the Nikkei 225. Finally, I expand the universe of stocks to include all liquid traded stocks in Japan and study the cross-sectional determinants of their comovement with stocks inside the Nikkei 225 index. Controlling for index membership, index overweighting is a significant determinant of the comovement of returns with index returns.

The cross-sectional evidence in this paper can be used to distinguish between two non-fundamentals based views of excess comovement, articulated by Barberis, Shleifer, and Wurgler (2004). In the first view, also adopted in this paper, comovement of stock returns comes from correlated investor demand shocks for a particular group of securities.[2] In the second view, dubbed the "information diffusion view", changes in short-term comovement come from differences in the speed at which security prices reflect new information. In the context of the Nikkei 225 stock index, this theory says that index members should incorporate information about aggregate earnings today, while non-index stocks incorporate this information with a lag. Both the demand theory and the information diffusion theory are consistent with index additions (deletions) experiencing increases (decreases) in short-term comovement with other index stocks.

To distinguish between the two theories, Barberis, Wurgler and Shleifer (2004) and Greenwood and Sosner (2004) examine changes in the cross-autocorrelation of index and non-index stocks following additions and deletions. The information diffusion theory predicts

[2] Barberis, Wurgler, and Shleifer (2004) further break down this first theory into an "investor habitat" view and a "category" view. Since both have similar theoretical implications, I treat them as subsets of a demand based theory of return comovement.

reductions in autocorrelation and cross-serial correlation with the index return following index addition, with converse results for deletions. The time series implications of the demand view, however, depend on the persistence of investor demand and the speed at which arbitrage is able to bring prices back to fundamentals. In any case, there appears to be evidence to support both theories. This ambiguity is eliminated in this paper, because the information diffusion theory does not have clear-cut cross-sectional predictions. While one can argue that index addition or deletion changes the speed at which a stock incorporates new information, there is no reason that after controlling for index membership, that weighting should be related to the speed at which a stock incorporates new information in the cross-section. Perhaps a nuanced theory of information diffusion would hold that even within a stock index, some stocks incorporate information faster than others do. But it seems unreasonable that this should depend on the price at which a stock entered the index many years ago. In short, the results favor the interpretation that excess comovement of stock returns comes primarily from behavioral, rather than informational, inefficiencies.

The distinction between the demand view and the information diffusion view of comovement is of both theoretical and practical importance. In the latter theory, some security prices do not commove enough, while in the former theory, some security prices commove too much. In the information diffusion theory, index membership increases the overall efficiency of the pricing process. In the investor demand theory, supported by the evidence, index members are subject to frequent and economically significant mispricing.

The magnitude of the results suggests a fundamental flaw in arbitrarily weighted stock indexes, and may explain their decline in popularity among sophisticated investors (Forbes, 1994 and Reuters, 2000). Recently, several sets of broadly used stock indexes have moved towards "float adjusted" weightings, in which index weight is based on the tradable capitalization of each stock.[3] In Japan, even the value weighted Tokyo Stock Exchange index (TOPIX) recently

[3] Recent conversions to float weighted indices include the MSCI global indices, the FTSE (United Kingdom), certain S&P indices, STOXX (Global), and SENSEXX (India).

announced a move towards free-float adjusted indices, citing a desire to "avoid supply and demand distortion of share prices" arising in the trading of index shares from "passive funds."[4] This motivation is consistent with the economically significant distortions I document.

The paper proceeds as follows. Section I develops predictions from a cross-sectional model of excess comovement. Section II describes the data and Nikkei 225 index methodology. Section III examines cross-sectional determinants of changes in comovement for a limited sample of Nikkei 225 additions and deletions. Section IV examines determinants *of the level* of comovement with index and non-index stocks, using a large sample of stocks traded in Japan. Section V concludes.

I. Security demand and comovement of stock returns: cross-sectional predictions

This section outlines a set of cross-sectional predictions concerning the relation between index overweighting and the comovement of stock returns. These predictions come out of a simple limits-to-arbitrage model in which uninformed investor demand shocks occur in proportion to a weighting vector. Although the *level* of index demand in any period is random, the proportionality between index demand shocks generates a cross sectional relation between index weights and the comovement of stock returns.

The predictions, although novel, can be generated within a theoretical framework that has been developed in other papers (e.g. Hong and Stein, 1999; Barberis, Shleifer and Wurgler, 2004; Greenwood and Sosner, 2004). I therefore present the model in reduced form.

In both the model and its empirical implementation, I measure comovement as the covariance between a stock's return and the equal weighted return of the other stocks in the index. Because the actual index is not equal weighted (thus generating the variation in demand shocks), the equal weighted return differs from the index return. I use the equal weighted return for two reasons. First, covariance with the equal weighted return provides a measure of the comovement with all index stocks. Second, I avoid misattributing the results to a mechanical

[4] http://www.tse.or.jp/english/news/2004/200407/040723_a.html , 2004/7/23, Press release

relation between index weight and comovement that comes from overweighted stocks contributing more to index return. For the purposes of notation, "index return" refers to the actual return of the Nikkei 225 index, while "equal weighted index return" refers to the equal weighted return of the 225 members of the Nikkei index. These two time series are 93% correlated during the sample period, and the main results go through with either measure.

A. Setup

The capital market contains a single risk-free security and N risky securities paying uncertain liquidating dividends $D_{i,T}$ in period T. Each risky security is in fixed identical supply Q. M ($M<N$) securities are part of a stock index. The risk-free asset is in perfectly elastic supply with a net return normalized to zero. The information flow for the dividend is given by

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

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 the two-factor model

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

$f_{M,t}$ and $f_{i,t}$ are the market and the idiosyncratic factors, respectively. All factors are normally distributed with zero mean and variance of one and are independently and identically distributed over time. The idiosyncratic factors $f_{i,t}$ are orthogonal to the market factor and are uncorrelated among securities. From (2), the covariance structure of the information process

$$\left(\Sigma_D\right)_{ij} \equiv \operatorname{cov}\left(\varepsilon_{i,t}, \varepsilon_{j,t}\right) = \begin{cases} 1, & \text{if } i = j \\ \psi_M^2, & \text{if } i \neq j \end{cases} \tag{3}$$

B. Trading Behavior

Two types of agents operate in the capital market – index traders and arbitrageurs. Arbitrageurs are risk-averse myopic investors with exponential utility of wealth. Index traders transfer funds into and out of the N stocks based upon their sentiment level u_t, and proportional to a weighting vector k. Thus,

$$u_{i,t} = k_i u_t \tag{4}$$

Where u_t is independently distributed over time and normal with mean zero and standard deviation σ_u^2. The structure of index demand is as follows

$$\begin{aligned} k_i &\geq 0, \qquad 1 \leq i \leq M \\ k_i &= -1, \qquad i > M \end{aligned} \tag{5}$$

If demand is positive for one index stock, then it must be positive for all index stocks, in proportion to index weight. If index demand is positive for index stocks, then it must be negative for all stocks outside of the index. Intuitively, this means that index investors pull funds out of other stocks to purchase the index in proportion to index weight.[5] For simplicity, I restrict demand for all non-index stocks to be identical.

Finally, weights of index stocks are normalized so that the sum of index weights is equal to one. This implies

$$\sum_{i=1}^{M} k_i = 1 \tag{6}$$

This normalization allows for a direct mapping from the model to the empirical work.

C. Predictions

Capital market equilibrium is obtained through the market clearing of security demands of index traders and arbitrageurs. Because index trader demand in each period is exogenous,

---

[5] Another specification might require $k_i=0$ for $M<i\leq N$. But this has no cross-sectional predictions for the relation between index weight and comovement between index and non-index stocks. Since the data conform with the modeling choice, I do not revisit the issue.

price levels are determined by the willingness of arbitrageurs to absorb demand. To obtain an analytical solution for returns, I assume that the economy is in a covariance stationary equilibrium, in which arbitrageurs correctly conjecture the conditional covariance matrix of future stock returns. Under these conditions, returns are a linear function of fundamental news and the index demand shock. In reduced form, asset returns can be expressed as

$$\Delta P_{i,t} = \varepsilon_{i,t} + \phi k_i u_t \tag{7}$$

where ϕ is a constant and $\Delta P_{i,t}$ is the change in price between $t-1$ and t. Henceforth, I will refer to $\Delta P_{i,t}$ as the return.[6]

I next define the equal weighted index return as the average return of the index constituents 1 through M.

$$\Delta P_{Index,t} = \frac{1}{M}\sum_{i=1}^{M}\Delta P_{i,t} \tag{8}$$

Note that the equal weighted index return differs from the actual index return, in which securities have different weights. The covariance between the returns of security j and the equal weighted index return is given by

$$\operatorname{cov}(\Delta P_{j,t}, \Delta P_{Index,t}) = \operatorname{cov}(\varepsilon_{j,t} + \phi k_j \Delta u_t, \frac{1}{M}\sum_{i=1}^{M}(\varepsilon_{i,t} + \phi k_i \Delta u_t)) = \psi_M^2 + k_j \phi^2 \sigma_u^2 \tag{9}$$

where k_j is positive for index securities. This leads to the first hypothesis.

Hypothesis 1. For an index security j, the covariance of returns with the returns of other index securities is increasing in index weighting k_j.

I next compute the covariance between the returns of index security i and a (representative) non-index security j

$$\operatorname{cov}(\Delta P_{i,t}, \Delta P_{j,t}) = \operatorname{cov}(\varepsilon_{i,t} + \phi k_i \Delta u_t, \varepsilon_{j,t} - \phi \Delta u_t) = \psi_M^2 - k_i \phi^2 \sigma_u^2 \tag{10}$$

[6] Greenwood (2004) derives the mapping between price changes used in models such as these and the units used in empirical work.

This leads to the second hypothesis

Hypothesis 2. The covariance of returns between index security i and non-index returns is decreasing in the index weighting of i.

Noting that k_j is equal to -1 for non-index stocks, we can analyze the change in comovement following addition or deletion from the index. Substituting $k_j = -1$ into equation (9) and contrasting with the result from *Hypothesis 1* yields the next hypothesis.

Hypothesis 3. Upon inclusion into a stock index, additions should experience increases in comovement with index stocks. The amount of increase in comovement is increasing in the weight at which the addition enters the stock index. Conversely, upon deletion from a stock index, deletions should experience decreases in their comovement with other index stocks. The amount of decrease in comovement is increasing in the weight the deletion held in the index prior to deletion.

A similar substitution in equation (10) yields the final hypothesis.

Hypothesis 4. Upon inclusion into a stock index, additions should experience decreases in comovement with non- index stocks. The amount of decrease in comovement is increasing in the weight at which the addition enters the stock index. Conversely, upon deletion from a stock index, deletions should experience increases in their comovement with non-index stocks. The amount of increase in comovement is increasing in the weight the deletion held in the index prior to deletion.

D. Discussion and Empirical Implementation

The basic empirical strategy is to confirm the cross-sectional relations suggested by Hypotheses 1 and 2. Using a full panel of stock returns in Japan, I create measures of comovement of stock returns with index returns, all estimated using past data. I then look at the cross-sectional relation between these comovement measures and overweighting of the stock in the Nikkei 225 index. Following the model, comovement should be estimated using a surrogate for the Nikkei 225 return, in which stock returns are equal weighted.

The model describes the covariance between security returns and the equal weighted index return. Because the covariance of fundamentals is the same between any two securities, all of the hypotheses developed above can also be read as statements about the *correlation* between security returns, not simply the covariance. I can take the predictions about correlation to the data, as long as I control for cross-sectional difference in the exposure to fundamentals.

In the model, any index security with a weight exceeding *1/M* is overweighted. The feature of the model that ensures that these securities also commove more with index returns is that all securities are identical and present in equal supply. In practice, however, overweighting should be measured relative to a security's ability to absorb demand without a large change in price. Although I experiment with several measures of overweighting, the one I use in most of the empirical tests is the log ratio of the index weight to the weight the stock takes in a value weighted index.

It may seem that *Hypothesis 3* and *Hypothesis 4*, which focus on changes in comovement, offer cleaner tests of the model than the determinants of comovement *levels* suggested by Hypothesis 1 and 2. In fact, an analysis of changes in comovement around index changes runs into a number of problems. First, because additions and deletions occur during different points in time, changes in comovement may reflect the changing variance of shocks to index demand (σ_u^2 in the model). Thus, stocks that experience particularly large drops in index beta upon index deletion may have been overweighted, or they may have been part of the index during a period of high variance in index trader demand. Second, there are very few additions and deletions for which I can collect a sufficient amount of returns data both pre- and post-event, subjecting these tests to a lack of power. However, drawbacks in mind, I examine the evidence for *Hypothesis 3* and *Hypothesis 4* in Section III.

II. Data

This section describes the data used in the study. The first part outlines sample construction. The second part discusses index methodology and the calculation of overweighting.

A. Sample Construction

The main sample consists of 298 stocks that were present in the index for at least 200 days between September 1, 1993 and August 29, 2003. This period is chosen because stock return and volume data is available for each stock using Datastream. Prior to 1993, I am unable to collect comprehensive data on returns for all index stocks.

Table I outlines the composition of this sample. Of the 298 stocks, 225 are in the stock index at any point in time.[7] Out of 225, there are 164 stocks in the index for the entire sample period. 73 stocks were added to the index between 1993 and September 2003. These stocks are matched by 73 deleted stocks. Of the 73 deleted stocks, however, only 41 provide enough data to be tracked before and after deletion. The stocks that do not have data following deletion were removed from the index due to delistings, mergers, and acquisitions.

Panel A of Table I shows the cross-sectional mean, standard deviation, and extreme values of time series averages of selected descriptors of the stocks in the sample. On average, each of the 298 stocks was in the Nikkei 225 stock index for 76 % of the sample. Index stocks experience a moderate amount of trading, with an average of 1% of shares trading daily.[8] The table also reports data on various characteristics used as controls in the cross sectional regressions. The average stock has a market capitalization of about ¥ 910 billion (approximately US$ 7.6 billion at time of writing), a price-to-book ratio (share price divided by book value per share) of 2.39, and a leverage ratio (long-term debt to common equity) of about 1.8.

[7] During the entire 10 year period, there are a total of 39 days during when less than 225 stocks were present in the Nikkei 225 index. These periods typically occur following an unexpected deletion. These periods are 3/27/2001 – 4/2/2001, 12/5/2001 – 12/11/2001, 2/25/2002 – 2/26/2002, 3/4/2002 – 3/5/2002, and 3/26/2002 – 4//1/2002.

[8] This figure is much lower if I remove a few outliers from the volume data.

Although not part of the main sample, I collect data on 1458 other stocks that were not in the index at *any* point between 1993 and 2003. Although most of these stocks are smaller than the median index constituent, some are much larger. Panel B of Table I shows that these stocks have lower average returns, trade less frequently, and are somewhat smaller than the index constituents. Price-to-book ratios for the two samples appear similar, while non-index stocks have lower leverage.

B. Nikkei 225 index methodology and overweighting

The value of the Nikkei 225 ($P_{N225,t}$) is determined by adding the ex-rights prices ($P_{i,t}$) of its constituents, divided by the face value (FV_i), times a constant, dividing the total by the index divisor (D_t):

$$P_{N225,t} = \frac{\sum_{i=1}^{225} \frac{P_{i,t}}{FV_i / 50}}{D_t} \qquad (11)$$

Most stocks have a face value of 50, though some have face values of 500, 5000 or 50000. The index divisor is adjusted daily to account for stock splits, capital changes, or stock repurchases. It is designed to preserve continuity in the index, though not necessarily in the index weights of its constituents. For example, following a two-for-one stock split of an index constituent, the effective weight of the stock falls by half, while the divisor is changed to keep the Nikkei index value unchanged.

After adjusting for face value, the index value is equal weighted in prices. This means that the index return, denoted R_{N225}, is the price-weighted average of the returns of its constituents

$$R_{N225,t} = \sum_{i=1}^{225} w_{N225it} R_{it} \qquad (12)$$

where index weights, w_{it}, are given by

$$w_{N225,it} = \frac{\frac{P_{it}}{FV_i/50}}{\sum_{j=1}^{225} \frac{P_{jt}}{FV_j/50}} \tag{13}$$

$w_{N225,it}$ can be interpreted as the cash value of stock i held by an investor at time t who owns one yen worth of the index. Denoting the yen amount of an index demand shock by K_t, $w_{it}K_t$ is the cash amount of stock i purchased by index traders.

Table 2 describes index weights between 1993 and 2003 in detail. Because index composition varies over time, a single cross-section would leave out 73 of the 298 securities. Therefore, for the purposes of the table, I compute index weights for each stock at the time when the stock enters the index, or on September 1, 1993, whichever comes later. The table shows that there are 298 stocks in the full sample. Most of these stocks have a face value of 50, but four have a face value of 500, two have a face value of 5000, and twelve have a face value of 50000.

Starting with the stocks of face value 50, the average price is 1,491, yielding an average Nikkei 225 weight of 0.58 percent.[9] As a measure of overweighting, I calculate the ratio of each stock's weight in the Nikkei to its weight in the market value weighted index

$$\frac{w_{N225,it}}{w_{VW,it}} = \frac{\dfrac{\frac{P_{it}}{FV_i/50}}{\sum_{j=1}^{225} \frac{P_{jt}}{FV_j/50}}}{\dfrac{MV_{it}}{\sum_{k=1}^{N^*} MV_{kt}}} \tag{14}$$

[9] Note that 0.58 x 225 > 100%. This occurs because additions tend to have high weights entering the index.

where $w_{VW,it}$ denotes the weight of stock i in the value weighted index, MV_{it} denotes the market value of stock i and N^* denotes the total number of tradable stocks on Section 1 of the Tokyo Stock Exchange. The value weighted TOPIX index is proportional to market capitalization of Section 1 stocks, so the overweighting measure in (14) can be calculated by dividing Nikkei weights by TOPIX index weights. The table shows that on average, the Nikkei weight exceeds the TOPIX weight by a factor of 5.7. The average ratio of Nikkei weight to market value weight can exceed one for two reasons. First, the Nikkei index includes less than 100% of the traded securities in Japan. Second, smaller stocks tend to be overweighted in the Nikkei, but receive equal weighting in reported averages.

After breaking down the full sample by face value, Table II shows average index weights, and overweighting, for the 73 additions, 73 deletions, and 164 remainders. The ratio of index weight to market value weight of the additions is low relative to the full sample average. On the other hand, the ratio of index weight to market value of the deletions is extremely high (an average of over 10). This may be why these stocks are deleted in the first place – they exert undue influence on Nikkei returns, relative to their weight in the market. It may also explain why the declines in comovement experienced by index deletions tend to be larger than the increases experienced by the additions.

Ultimately, the quantity of interest is not the *average* over- or underweighting of stocks in the Nikkei 225, but the cross sectional differences in weighting across stocks. Panel A of Figure 1 shows the histogram of index weights, using data from the 225 index constituents at the (exact) mid point of the sample. The distribution is heavily left skewed, with most stocks taking index weights between 0 and 0.15%. Panel B shows that the skewness of the weight distribution is reduced substantially when I scale the weights by market value weight, following equation (14). However, 28 stocks are still underweighted relative to the TOPIX.

I define overweighting, *OW*, as the log of one plus the ratio of Nikkei weight to TOPIX index weight (presented in Panel B of Figure 1)

$$OW_{it} = \log\left(1 + \frac{w_{N225,it}}{w_{MV,it}}\right). \tag{15}$$

The distribution of overweighting for the 225 stocks in the Nikkei 225 index at the mid point of the sample is shown in Panel C. Because Nikkei 225 weight is 0 for non-index stocks, overweighting is equal to zero for these stocks (log (1+0)). The last column of Table II summarizes overweighting for the entire sample of 298 securities.

For the remainder of the paper, 'overweighting' refers to the quantity computed in equation (15). Note that in each cross-section, the overweighting vector OW_t, is an instrument for the true weighting vector (k in the model), because actual demand shocks in each period are unobserved.

III. Security demand and changes in comovement for index additions and deletions

This section briefly analyzes cross sectional determinants of changes in comovement for stocks entering and leaving the Nikkei 225 index. According to the theory, additions should experience increases in comovement with other index stocks. *Hypothesis 3* suggests that the amount of increase should be proportional to overweighting of the addition at time of entry. Additions should also experience decreases in comovement with non-index stocks, with the amount of decrease proportional to overweighting at time of entry. Conversely, deletions should experience opposite effects on removal from the index, proportional to the overweighting at time of exit.

Between 1993 and 2003, 73 stocks were deleted from the Nikkei index. Many of these deletions occurred because of mergers, acquisitions, or delistings, while some were at the discretion of the index organizers. To track the comovement of the additions and deletions with other index stocks, I estimate rolling univariate time series regressions of stock returns on the equal weighted return of the 168 stocks that are in the Nikkei index for the entire sample period

$$R_{it} = \alpha_{it} + \beta_{EWN225,it} R_{EWN225,t} + \varepsilon_{it} \qquad (16)$$

where $R_{EWN225t}$ is the equal weighted Nikkei return. This regression is estimated for every period t and every security i, using 100 days of past returns. Note that both α and β are subscripted with *it*. This is to indicate that the coefficients are specific to a security and a period [t-100,t]. Figure 2 plots the average β_{EWN225} coefficients from this regression, plotted in event time and separately for the additions and deletions. The figure confirms increases in comovement following inclusion into a stock index, and decreases in comovement following deletion from the index (Barberis, Shleifer, and Wurgler, 2004; Greenwood and Sosner, 2004; and Coakley and Kougolis, 2004). Panel B also plots the average R^2 from this regression, plotted in event time.

To analyze the determinants of the change in comovement, I compute $\Delta\beta_{EWN225}$, the change in the slope parameter from (1) around addition or deletion from the index. Unfortunately, due to post-event data availability, I am only able to calculate the change in comovement for 61 out of the 73 additions, and for 39 of the deletions. Panel A of Figure 3 plots $\Delta\beta_{EWN225}$ against the overweighting of the stock in the index. To facilitate a comparison of the additions and deletions, the y-axis plots $\Delta\beta_{EWN225}$ for the additions, and $-\Delta\beta_{EWN225}$ for the deletions. For the additions, the overweighting is measured on the day after the stock is added to the Nikkei 225. For the deletions, overweighting is measured on the last day that the stock is a member of the index.

The figure shows that in spite of data limitations, there is a discernible correlation between index overweighting and the change in comovement. To test this cross-sectional relation more formally, Panel A of Table III shows the results from a cross-sectional regression of $\Delta\beta_{EWN225}$ on index overweighting, shown separately for the additions and deletions. The table confirms a significant relationship between overweighting and the change in comovement following index changes. On the right-hand-side columns of the table, I run the joint regression that includes the additions and deletions. For this estimation, I reverse the sign on the change in comovement for the deletions, to make the two samples comparable. The table shows that for the combined sample, the relation between over weighting and the change in comovement is stronger.

Hypothesis 4 suggests a relation between between index overweighting and changes in comovement between index stocks and non-index stocks. To test this, I estimate rolling bivariate time series regressions of stock returns on the equal weighted return of the remainder stocks, and the return on the value weighted TOPIX

$$R_{it} = \alpha_{it} + \beta^{*}_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it} \tag{17}$$

collecting regression coefficients β^{*}_{EWN225} and β_{TOPIX}. The asterisk on the beta on the equal weighted Nikkei return is to distinguish it from the same coefficient in the univariate regressions. I compute changes in comovement with index stocks, $\Delta\beta^{*}_{EWN225}$, and with non-index stocks, $\Delta\beta_{TOPIX}$, following each addition and deletion. Panel B of Figure 3 plots $\Delta\beta^{*}_{EWN225}$ against the overweighting of the stock in the index, and Panel C plots $\Delta\beta_{TOPIX}$ against the overweighting. Following the same procedure as before, the y-axis plots $\Delta\beta$ for the additions, and $-\Delta\beta$ for the deletions.

The figures show a clear positive relation between index overweighting and the change in comovement with other index stocks, and a clear negative relation between index overweighting and the comovement with stocks outside of the index. To test this cross-sectional relation more formally, Panel A of Table III shows the results from cross-sectional regressions of $\Delta\beta_{N225}$ and $\Delta\beta_{TOPIX}$ on index overweighting. There is a significant, but weak relationship between over weighting and conditional changes in comovement, as suggested by Hypothesis 4. Results for the additions are weaker. These results are consistent with the hypothesis, laid out in the model, that index investors pull funds out of non-index stocks to buy the Nikkei 225 stocks, in proportion to index weight.

Unfortunately, the tests that focus solely on the small sample of additions and deletions suffer from the drawbacks discussed in Section I. In particular, the tests are not truly cross-sectional as the additions and deletions occur at different points in time. More importantly, these tests eliminate the entire cross sectional variation between index weights of stocks that remain in the index. This in mind, the next section turns to the full panel.

IV. Security demand and the cross section of comovement

This section analyzes the cross sectional determinants of the comovement among index stock returns and between index and non-index stock returns. I begin by visually examining the determinants of comovement for a single cross section of stocks in the middle of the sample. I then do a more systematic analysis. This is done in two parts. First, I study the cross-sectional determinants of comovement for stocks that are in the Nikkei index, or were once present in the index. Then, I expand my approach to analyze the cross-sectional determinants of comovement for all stocks traded on Section I of the Tokyo Stock Exchange. Finally, I subject all of the results to a series of robustness tests.

A. Graphical analysis of comovement for a single cross section

I start by studying the determinants of index comovement for the 225 stocks present in the Nikkei 225 index on September 1, 1998, the mid point of my sample. For each of the 225 stocks, I estimate univariate time series regressions of stock returns on the equal weighted return of the remainder stocks, following equation (16). The slope parameter β_{EWN225} is a measure of the comovement of the stock return during the previous 100 days with the equal weighted index return. An alternate measure of comovement, which I also use, is the R^2 from this regression.

Panel A of Figure 4 plots equal weighted index beta, β_{EWN225}, against the lagged overweighting, *OW*. In the figure, overweighting is defined as the log of 1 plus the ratio of the Nikkei 225 weight to the TOPIX weight, and is measured 100 days earlier, on April 14, 1998. The timing is designed to ensure that the cross sectional relation is not driven by changes in prices, and hence changes in index weights, during the period in which beta is estimated.

The figure shows a strong positive relation between index overweighting and index beta. In the cross-sectional regression that corresponds to this figure (unreported), the slope is 0.28, and the R^2 is 0.25. Combined with a cross-sectional standard deviation of index overweighting of 0.78, the figure shows that a one standard deviation increase in overweighting is associated with a change in index beta of 0.22.

Panel B plots R^2 against lagged overweighting. Again, the figure shows a strong positive correlation between these two variables. This shows that the cross sectional relation is driven by increased correlation of returns with index returns for overweighted stocks, rather than by their increased variance.

I next study the relation between index overweighting and the *conditional* comovement of stock returns with stocks inside and outside of the index. Using the same cross section, I estimate bivariate time series regressions of stock returns on the equal weighted return of the 164 remainder stocks, and the return on the value weighted TOPIX index, following equation (17). The slope parameter β^*_{EWN225} is a measure of the conditional comovement of the stock return during the previous 100 days with the equal weighted index return. The slope parameter β_{TOPIX} is a measure of the conditional comovement with stocks outside of the index.

Panel A of Figure 5 plots β^*_{EWN225} against lagged overweighting. Consistent with the previous results, the figure shows a strong positive relation between these two variables. The cross-sectional R^2 from this regression is 0.59.

Panel B of Figure 5 plots β_{TOPIX} against lagged overweighting. Confirming *Hypothesis 2*, the figure shows a strong negative relation between index overweighting and the conditional comovement of index stocks with stocks outside of the index. The strength of the result is consistent with index investors pulling funds out of other stocks in order to invest in index stocks, in proportion to index weight.

B. Panel Analysis: Comovement among Nikkei 225 stocks

I now modify the cross-sectional approach from above to study comovement for the full panel of stock returns between 1993 and 2003. The basic approach is as follows. Every 100 days starting on January 20, 1994 (100 days after the first day of returns), I estimate time series regressions of returns on the equal weighted index return for the past 100 days, following equation (15). I estimate these regressions for every security in the sample, provided they have complete returns data over the previous 100 days, and provided their index status does not change during the estimation period. The full ten-year panel contains 26 cross-sections, with an average of 275 stocks in each. For each security and each cross section, I record the beta from this regression, as well as the associated R^2.

The next step is to relate these measures of comovement in each cross section TO the index overweighting at the start of the period. To do this, I run 26 cross sectional regressions of the comovement measures from (15) on index overweighting at the start of the period

$$\beta_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + u_{it}, \tag{18}$$

and

$$R^2_{it} = a_t + b_t \cdot OW_{it-1} + u_{it} \tag{19}$$

The first line of Table IV shows the time series average of coefficients *a* and *b* from (18), as well as their associated t-statistics, following Fama and Macbeth (1973). The average sensitivity of index beta to overweighting is a highly significant 0.218. Note that in *all* 26 cross-sectional regressions, the coefficient *b* is positive (unreported). Moreover, the table indicates that the time series average of R^2 from this cross-sectional regression is 0.213. Put simply, index overweighting explains, on average, more than twenty percent of the comovement among index stocks.

Line 2 of Table IV shows the time series average of coefficients from the cross-sectional regressions of R^2 on lagged index weight. Similar to the previous results, index overweighting explains a significant fraction of the correlation between the returns of index stocks.

I next alter the second stage regressions to include controls for other common factors in returns. Banz (1981), Fama and French (1992), find that commonality in average returns can be attributed to size and the book-to-market ratio. I thus run a second set of regressions that control for the (log of) size and the price-to-book ratio of each stock

$$\beta_{EWN225,it} = a_t + b_t \cdot X_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it}, \tag{20}$$

and

$$R^2_{it} = a_t + b_t \cdot X_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it} \tag{21}$$

Where *P/B* is the price to book ratio (share price divided by the book value per share for the appropriate financial year end, adjusted for capital changes) and *Size* is the log of market capitalization. These results are shown in lines 3 and 4 of Table IV. Comovement of a stock with other index members is negatively related to the price-to-book ratio, and negatively related to log size. The coefficient *b* on overweighting falls slightly to 0.161, but remains significant. The cross-sectional results on R^2 also remain strong.

There is a possibility that some of these results are driven by the fact that in each cross-section, there are an average of 50 stocks (275 minus 225 Nikkei members) that have a Nikkei 225 index weight of zero. Thus, a simple difference in comovement averages among Nikkei stocks relative to stocks outside of the Nikkei could generate an empirical relation between index weight and comovement. To lay this concern to rest, I repeat the second stage regressions, restricting each cross section to contain only stocks that remained in the index between *t-100* and *t*, where *t* denotes the time of measurement. In Table IV, these results are designated by setting the sample equal to "index members." The results for this subset of stocks remain strong.

Finally, I repeat all of these tests using weekly returns data. This reduces the total number of cross-sections to 10, increasing in the standard errors. *A priori*, at longer horizons, one would expect excess comovement to weaken as arbitrage brings prices back to fundamentals, or as demand dissipates. The results for weekly returns are indeed weaker – the coefficient *b* on lagged overweighting is lower across the board. However, in 6 out of the 8 regressions, overweighting is still a strong and statistically determinant of comovement. In the remaining 2 regressions, the positive relation between comovement and overweighting is reversed.

C. Panel Analysis: Comovement with non-index stocks

Hypothesis 2 says that the covariance of returns between index security *i* and non-index returns is decreasing in the index weighting of *i*. I test this proposition using a two step procedure.

In the first stage, I generate repeated time specific cross-sectional estimates of conditional comovement of each stock with the equal weighted index return and with return of the value weighted TOPIX index. Every 100 days starting on January 20, 1994, I estimate bivariate time series regressions of returns on the equal weighted index return and the value weighted TOPIX return

$$R_{it} = \alpha_i + \beta^*_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it} \tag{22}$$

using 100 days of prior returns data. These regressions are estimated for every security in the panel, provided they have complete returns data over the estimation period, and provided their index status does not change during the estimation period. For each security and each cross section, I record β_{EWN225}, the conditional comovement of stock returns with the equal weighted index return, and β_{TOPIX}, the conditional comovement of stock returns with stocks outside of the index.

In the second stage, I relate these measures of comovement to lagged index overweighting and controls

$$\beta^*_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it} \tag{23}$$

These results are in Table V. Both the univariate and multivariate relations between conditional comovement and overweighting appear stronger than in the previous table. Line 3 and Line 4 show that when each cross section is constrained to include only stocks in the Nikkei index (omitting stocks with zero index weight), that the results are again stronger than in Table IV.

Panel B shows average coefficients from the regression of comovement with non-index stocks on overweighting and controls

$$\beta_{TOPIX,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it} \tag{24}$$

Line 1 shows that there is a strong negative relation between the comovement of a stock's returns with stocks outside of the Nikkei, and that stock's weight in the Nikkei index. Line 2 shows that this relation continues to hold after controlling for price-to-book and log size. Lines 3 and 4 show that this negative relationship remains strong and significant when each cross section is constrained to include only the stocks present in the Nikkei index during the preceding 100 days.

To summarize the results thus far, the empirical evidence is consistent with the claim that that index investors pull funds out of non-index stocks to buy the Nikkei 225 stocks, in proportion to index weight, and that this trading exerts a significant effect on security prices.

D. Expanded sample analysis

Some of the cross-sectional variation in index weights is driven by differences in the average index weights of index and non-index stocks. This is a concern because it could imply that index membership, rather than index weight, is driving the cross sectional results. Recall that if this were true, the results might support a wider range of interpretations, including one in which index membership changes the rate at which a stock incorporates new common information. This concern is somewhat alleviated in the regressions in which the sample is restricted to only index stocks (lines 3 and 4 of Table IV). In these specifications, all of the variation in index weights comes from index stocks. Casual inspection of Table IV suggests that because the slope coefficient *b* declines when the sample is restricted in this way, that a fraction of comovement may be explained by index membership.

In this section, I formally identify the distinct roles of index weighting and index membership. To do this, I expand the basic sample with an additional 1,458 stocks. With the expanded sample, I modify the second step of the two-step procedure to allow separate roles for index weighting and index membership in the cross section.

The first stage remains unchanged: I estimate time series regressions of returns on the equal weighted index return for the past 100 days, following equation (16), collecting measures of comovement in each instance. The second stage regression is modified to include a dummy variable for index membership, as follows

$$\beta_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + u_{it} \tag{25}$$

where *OW* denotes lagged index overweighting and the lagged indicator $1_{OW_{it-1}>0}$ takes a value of one if the stock was in the Nikkei 225 index, and zero otherwise. The other controls are defined as before. I estimate this regression for each cross section, reporting average coefficients in Table VI.

The table shows that even in the broader cross section, index overweighting has a profound effect on the comovement of returns. The slope in the univariate regression between index beta and overweighting averages 0.30. When I add the control for index membership, the average slope is reduced to 0.22. After adding the additional controls, the slope remains approximately the same, at 0.20.

While index overweighting retains a prominent role in the cross section, the regressions show that in addition, the dummy variable for index membership is occasionally significant. The simplest interpretation is that the significance of index membership is consistent with an information diffusion theory of excess comovement: index membership allows stocks to incorporate shocks to common information at a faster rate. However, there is a risk of overinterpretation, for two reasons. First, because the equal weighted index return incorporates all index stock returns, its comovement with index stocks is higher by design.[10] Second, index membership may proxy for other common characteristics of index stocks that cause them to have higher fundamental comovement. Of course, one could level the same omitted variable complaint against index overweighting *OW*. But it is hard to see why overweighting would be systematically correlated with *any* omitted characteristic.

Panel B of Table VI shows these same results estimated on weekly measures of the comovement of stock returns. The same general pattern emerges: both index membership and overweighting are significant determinants of comovement, although the multivariate results are weaker.

Finally, I apply this procedure to study the conditional comovement of returns with stocks outside of the index. For the expanded sample of stocks, I estimate bivariate time series

---

[10] This particular problem can be corrected by altering the first stage of the two step procedure, as follows. For each stock i in each period t, regress returns of i on the equal weighted return of all stocks in the index, excluding i. Because this has virtually no effect on the estimated β coefficients and significantly complicates the exposition, I have chosen the equal weighted index return instead.

regression of returns on the equal weighted index return and the value weighted TOPIX return, following (17). Then, in repeated cross-sectional regressions, I study the determinants of comovement with index and non-index stocks

$$\beta^*_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + u_{it} \tag{26}$$

and

$$\beta_{TOPIX,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + v_{it} \tag{27}$$

Similar to equation (25), the regressions now include a dummy variable that takes on a value of 1 if the stock was a member of the Nikkei index, and zero otherwise. Average coefficients from (26) and (27) are shown in Table VII. Panel A shows the determinants of comovement with Nikkei 225 stocks (β^*_{EWN225}) and Panel B shows the determinants of comovement with stocks outside of the Nikkei 225 (β_{TOPIX}).

These results are similar to those shown in Table VI, but stronger. Index overweighting has a strong positive relation with comovement of a stock with stocks in the index, and a strong negative relation with the comovement of a stock with stocks outside of the index. In the multivariate regressions (line 3 of Panel A and line 3 of Panel B), index membership does not enter significantly. In line 2 of Panel A, index membership enters with the wrong sign.

E. Robustness

The empirical analysis rests on a set of modeling choices. These concern (i) the length of time over which comovement is estimated, (ii) the calculation of standard errors, (iii) the measurement of index overweighting, (iv) the choice of control variables, and (v) calculation of index returns. This section examines whether altering any, or all, of those choices affects the results significantly.[11]

---

[11] The table with the results from this section is available from the author on request.

i. Length of time over which comovement is estimated

All of the reported results are based on betas estimated using 100 days (or 50 weeks) of daily (or weekly) return data. I expand the measurement window from 100 days to 250 days, and compute univariate equal weighted Nikkei 225 betas according to (16) and multivariate equal weighted Nikkei 225 and Topix betas according to (17). I then re-estimate the first two lines from Table V and Table VI. The total number of cross-sections is now reduced from 26 to 10, causing the standard errors to increase; the main results remain.

ii. Measurement of standard errors

In Tables IV, V, VI, and VII, t-statistics are based on the standard deviations of the cross-sectional regression estimates, following Fama and Macbeth (1973). This technique assumes that the time series of cross-sectional estimates is not autocorrelated. However, one can easily modify the procedure to incorporate autocorrelation in parameter estimates. I repeat my calculation of standard errors in Table IV and Table V, following Newey and West (1987) in my estimate of the spectral covariance matrix. Allowing autocorrelation up to two lags, the t-statistic on b falls from 10.17 to 7.28 (see Table IV for original). Allowing autocorrelation up to four lags, the t-statistic falls to 6.49.

iii. Measurement of index overweighting

The measure of index overweighting used throughout the paper is $log(1+w_{N225}/w_{TOPIX})$. As discussed previously, this seems like an intuitive measure of the overweighting of the stock . A similar measure of overweighting is the ratio of index weight to the weight the stock would have taken in the Nikkei 225, had the index been value weighted. Although this measure behaves slightly differently in the time series – as the ratio of Nikkei 225 market capitalization to TOPIX capitalization changes – it behaves identically in the cross section.[12]

---

[12] $\mathrm{Log}(1+w_{N225}/w_{N225MV})$ is a positive and monotonic transformation of $\mathrm{Log}(1+ w_{N225}/w_{TOPIX})$, because w_{TOPIX}/w_{N225MV} is constant in any period t.

An alternative scaling factor would be a more direct measure of a stock's liquidity. Ideally, this would capture a stock's ability to absorb uninformed demand without a change in price. Unfortunately, an off-the-shelf measure does not exist, although one may be able to construct price impact measure using trading data, which I do not have. Instead, I use the ratio of index weight to average daily trading volume, measured in yen. On average, this measure has a cross-sectional correlation with the index return of 81%. When I use this measure of overweighting in Table IV, the results go through as before.[13]

A final measure of overweighting is not to scale the index weights at all, setting overweighting equal to $log(1+w_{N225})$. This measure has two significant problems. First, running counter to intuition, it makes several liquid stocks with high market capitalization appear overweighted. Nevertheless, I experiment with it. The relation between conditional comovement and weighting in Table V continue to hold. The basic relation in Table IV, although still present, is no longer statistically significant.

iv. Choice of control variables

The control variables were selected based on two criteria. First, size and book-to-market are both extensively documented as having pronounced effects on the cross section of average returns. Second, both are available from Datastream for the majority of the stocks in my sample.

A third control variable, not included in the main tests, is firm leverage. It follows from Modigliani and Miller (1958) that controlling for asset risk, equity betas should be positively related to the degree of financial leverage. However, Hecht (2002) tests this proposition using U.S. data and finds that the data do not confirm it. Nevertheless, I collect a proxy for leverage from Datastream (see Table 1 for summary information), and repeat the basic multivariate test given in line 2 of Table IV. In Japan, leverage is strongly positively related to the comovement

[13] The volume data is not as reliable as the pricing data, with several stocks having trading days with reported turnover over 100%. Before I can construct the alternate overweighting measure described in the text, I winsorize the volume data in each cross-section at the 1 percent level.

of stock returns with the Nikkei. The coefficient on lagged overweighting, *b*, becomes both more economically and statistically significant following introduction of this control.

v. Calculation of equal weighted index return

All of the empirical tests measure index comovement using an equal weighted average of the returns of the 164 stocks that remained in the index for the entire sample period. This time series is 93 percent correlated with the actual Nikkei 225 return over the same period. It is also 99 percent correlated with the equal weighted return of the Nikkei, and 88 percent correlated with the value weighted return of index constituents.

Although these three alternatives lack a compelling theoretical motivation, I experiment replacing the actual Nikkei return for the equal weighted index return in (16) and (17). To start, I replicate the main results in Table IV using each of these series, with no change in the results. I also replicate the basic tests in Table V, with similar conclusions.

V. Conclusions

In the presence of limits to arbitrage, periodic investor demand shocks that vary in degree across securities should have cross-sectional effects on the comovement of their returns. This paper exploits the weighting system of the Nikkei 225 index in Japan to relate cross-sectional regularities in demand shocks to the comovement of stock returns. Using index overweighting as an instrument for the true variation in demand for Nikkei 225 stocks, I trace an empirical relation between overweighting and the comovement of a stock with other stocks in the index. Overweighting accounts for a large fraction of the cross sectional variation in comovement among index stocks, and comovement between index and non-index stocks. To summarize, the results show that proportionality between investor demand shocks can have powerful effects on the cross-section of stock returns.

I argue that the results are consistent with a demand theory of excess comovement, in which correlated investor demand for securities causes periodic and widespread mispricing. The results are not favorable for market friction based explanations of excess comovement, in which index membership is related to the speed at which stocks incorporate new common information.

The cross-sectional effects I document are likely to appear in stock indexes outside of Japan, and in broader settings in which there is variation in periodic demand shocks for stocks. Practically speaking, the magnitude of the results suggests that members of arbitrarily weighted stock indexes – oftentimes "liquid" securities – are subject to frequent mispricing. Under these circumstances, it is not hard to understand the declining influence of the price weighted Dow Jones Industrial Average (DJIA) relative to the value weighted S&P 500, or the growing influence of the TOPIX relative to the Nikkei 225.[14] Finally, my results may explain why several sets of broadly used stock indices have moved towards "float adjusted" weightings, in which index weight is based on the tradable capitalization of each stock.[15]

---

[14] See Forbes, 1994 and Reuters, 2000.

[15] Recent conversions to float weighted indices include the MSCI global indices, the FTSE (United Kingdom), certain Standard and Poors indices, STOXX (Global), and SENSEXX (India).

References

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161-199.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2004, Comovement, forthcoming *Journal of Financial Economics*.

Bodurtha, James, Dong-Soon. Kim and Charles M. Lee, 1995, Closed-end country funds and US market sentiment, *Review of Financial Studies* 8, 879-918.

Coakley, Jerry, and Periklis Kougoulis, 2004, Comovement and FTSE 100 index changes, Working Paper, University of Essex.

De Long, J.Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-38.

Fama, Eugene., and Kenneth. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 46, 427-466.

Fama, Eugene, and James MacBeth, 1973, Risk, Return and Equilibrium: Empirical Tests, *Journal of Political Economy*, Vol. 81, pp. 607-636.

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

Forbes, 1994, Inaccurate yardstick. (The Dow Jones Industrial Average). R.S. Salomon Jr, 11 April 1994.

Froot, Kenneth, and Tarun Ramadorai, 2003, Currency returns, intrinsic value, and institutional investor flows, forthcoming, *Journal of Finance*.

Greenwood, Robin, 2004, Short and Long Term demand curves for stocks: Theory and Evidence on the dynamics of arbitrage, forthcoming, *Journal of Financial Economics*.

Greenwood, Robin and Nathan Sosner, 2004, Trading patterns and excess comovement of stock returns, Working Paper, Harvard Business School.

Hecht, Peter, 2002, Do equity covariances reflect financial leverage?, Working Paper, Harvard Business School.

Hardouvelis, Gikas, Rafael La Porta, and Thierry Wizman, 1994, What moves the discount on country equity funds? in Jeffrey Frankel (ed.), The Internationalization of Equity Markets (Chicago, The University of Chicago Press).

Harris, Lawrence, 1989, S&P 500 Cash stock price volatilities, *Journal of Finance* 44, 1155-1175.

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

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

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-78.

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

Lee, Charles, Andrei Shleifer, and Richard 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* 15, 533-564.

Morck, Randall, Bernard. Yeung and Wayne Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous stock price movements?, *Journal of Financial Economics* 58, 215-260.

Newey, Whitney, and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Pindyck, Richard, and Julio Rotemberg, 1993, The comovement of stock prices, *Quarterly Journal of Economics* 108, 1073-1104.

Pindyck, Richard, and Julio Rotemberg, 1990, The excess comovement of commodity prices, *Economic Journal* 100, 1173-1189.

Reuters News, 2000. New Nikkei reflects Japan better, but flaws remain. R. Maeda, April 25.

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

Shleifer, Andrei, 2000, Inefficient Markets: An Introduction to Behavioral Finance (Oxford University Press, Oxford).

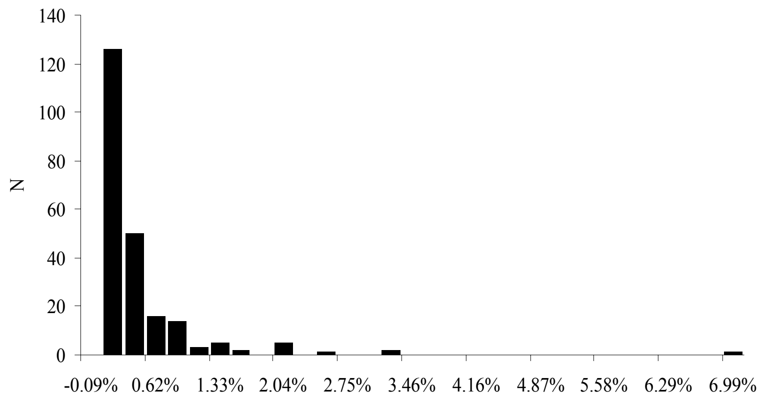
Vijh, Anand, 1994, S&P500 Trading strategies and stock betas, *Review of Financial Studies* 7, 215-251.

Figure 1

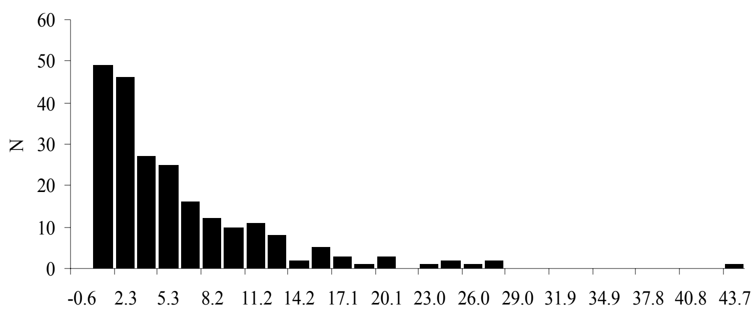
Nikkei 225 index weights and overweighting

Cross-sectional distribution of index weights of the stocks in the Nikkei 225 index on September 1, 1998, the mid-point of the sample. Panel A shows the distribution of raw index weights for the 225 stocks in this data. Panel B shows the ratio of index weights to the weight that each stock takes in the market value weighted TOPIX index. Panel C shows the distribution of the overweighting measure, equal to the log of one plus the ratio of index weights to the weight that each stock takes in the market value weighted TOPIX index.

Panel A. Nikkei 225 weights, mid-sample



Panel B. w_{N225}/w_{MV}, mid-sample



Panel C. Overweighting = $\log(1 + w_{N225}/w_{MV})$, mid-sample

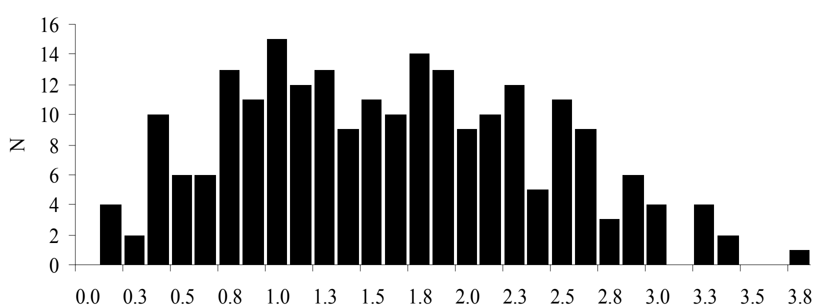
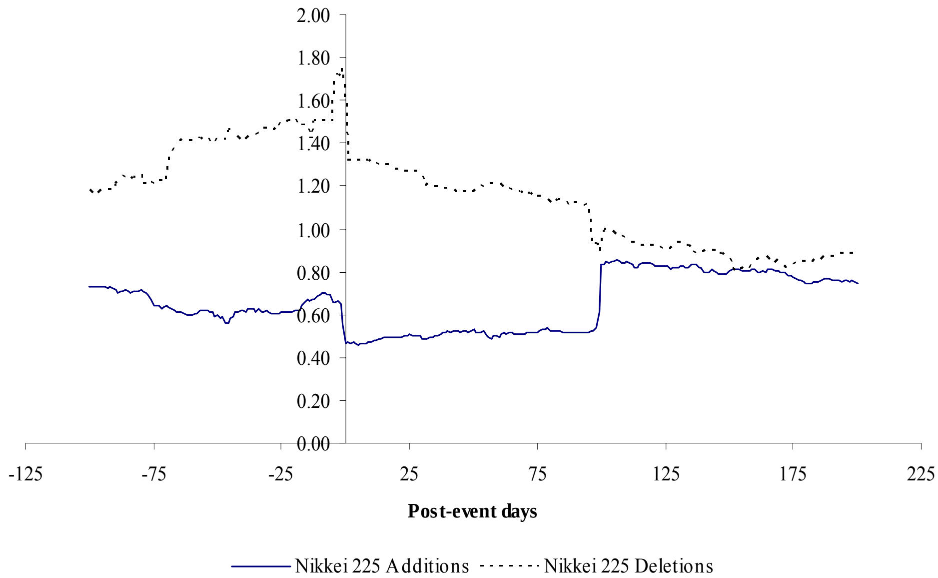


Figure 2

Comovement of additions and deletions before and after Nikkei 225 index changes, 1993-2003

Average slope coefficients and corresponding R^2 from simple rolling OLS regressions of daily returns of additions and deletions on the equal weighted daily return of the 164 stocks remaining in the Nikkei index between 1993 and 2003, plotted in event time. Additions (Deletions) include stocks added to (deleted from) the Nikkei 225 stock index between September 1993 and September 2003, which have complete trading data 200 days before and after addition or deletion. Panel A shows rolling average betas for additions and deletions, estimated using 100 days of backwards looking data. Panel B shows the corresponding rolling average R^2.

Panel A. Changes in Beta for Nikkei 225 additions and deletions



Panel B. Changes in R^2 for Nikkei 225 additions and deletions

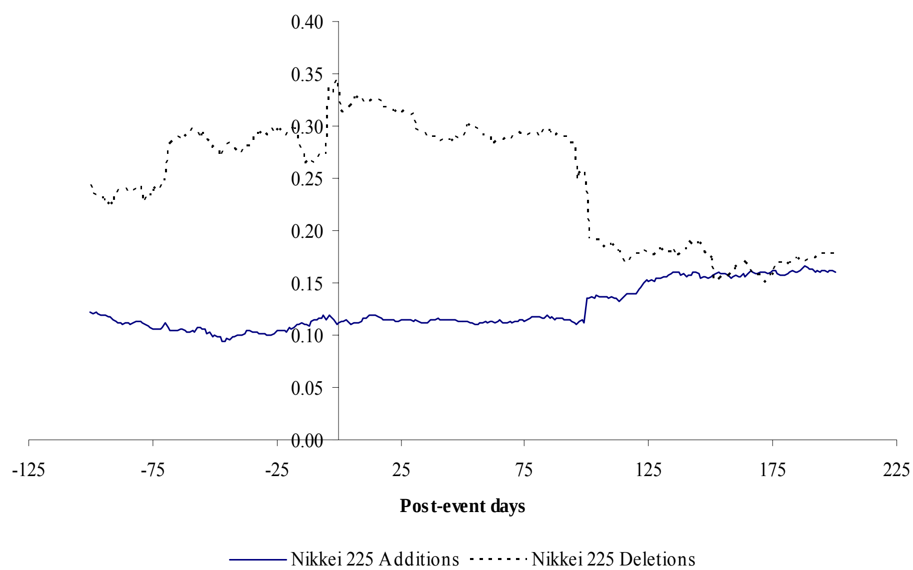


Figure 3

Changes in comovement and Nikkei 225 index overweighting

Changes in the comovement of stock returns with the returns of other index and non-index stocks, plotted against overweighting of the stock in the index. Index overweighting is equal to the log of one plus the ratio of index weight to the weight in the market value weighted TOPIX index. In Panel A, the change in comovement ($\Delta\beta_{EWN225}$) is the change in the slope coefficient from a regression of daily returns on the equal weighted return of the 164 stocks remaining in the Nikkei index between 1993 and 2003

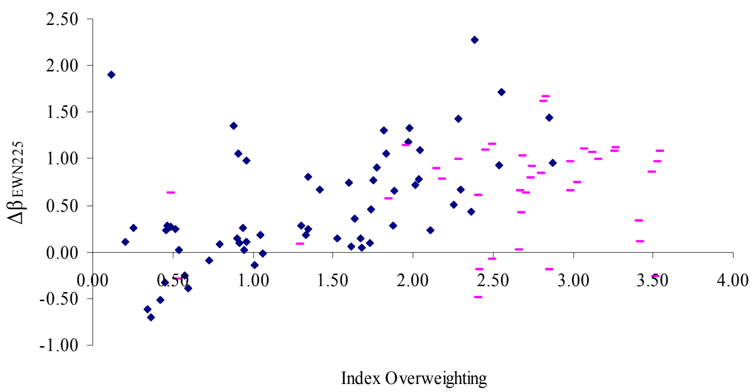
$$R_{it} = \alpha_i + \beta_{EWN225,it} \cdot R_{EWN225,t} + \varepsilon_{it}$$

In Panel B, the change in comovement ($\Delta\beta^*_{EWN225}$) is the change in the slope coefficient on the equal weighted Nikkei return, taken from a regression of daily returns on the equal weighted return of the stocks in the Nikkei 225 index, and the return on the value weighted TOPIX index

$$R_{it} = \alpha_i + \beta^*_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it}$$

In Panel C, the change in comovement ($\Delta\beta_{TOPIX}$) is the change in the coefficient on the TOPIX return from the above regression. It measures the conditional comovement with stocks outside of the Nikkei 225 index. All regressions are estimated using 100 days of prior return data. In each panel, the y-axis shows the change in comovement, the difference between comovement 120 days after index addition or deletion, and comovement 20 days before index addition or deletion. Additions (Deletions) include 61 (39) stocks added to (deleted from) the Nikkei 225 stock index between September 1993 and September 2003. Additions are marked with diamonds; deletions are marked with dashes.

Panel A. $\Delta\beta_{EWN225}$ (y-axis) vs. Overweighting (x-axis)



Panel B. $\Delta\beta^*_{EWN225}$ (multivariate, y-axis) vs. Overweighting (x-axis)

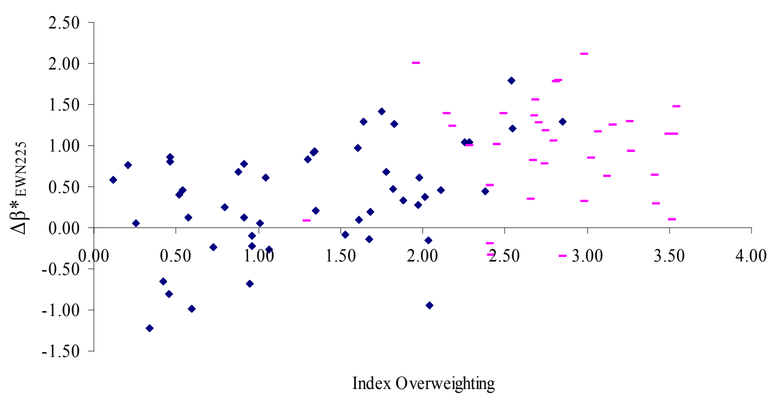


Figure 3 (Continued)

Changes in comovement and Nikkei 225 index overweighting

Panel C. $\Delta\beta_{TOPIX}$ (multivariate, y-axis) vs. Overweighting (x-axis)

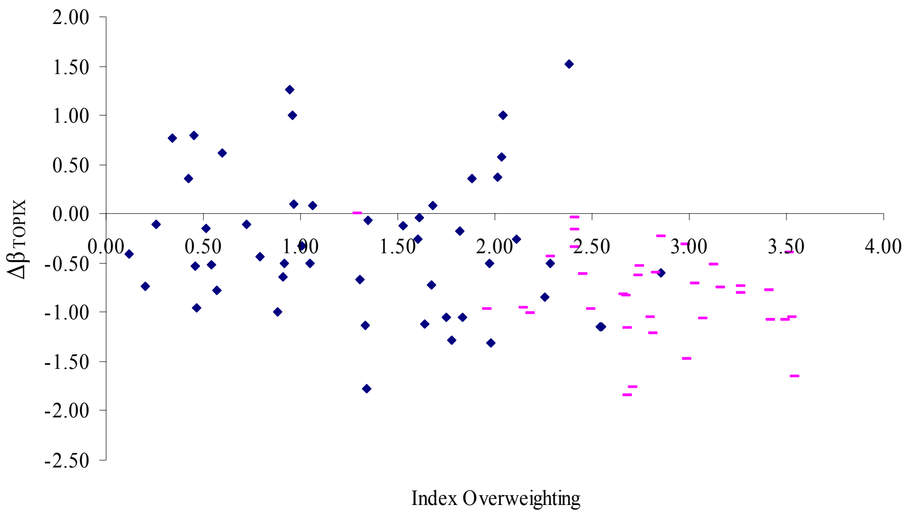
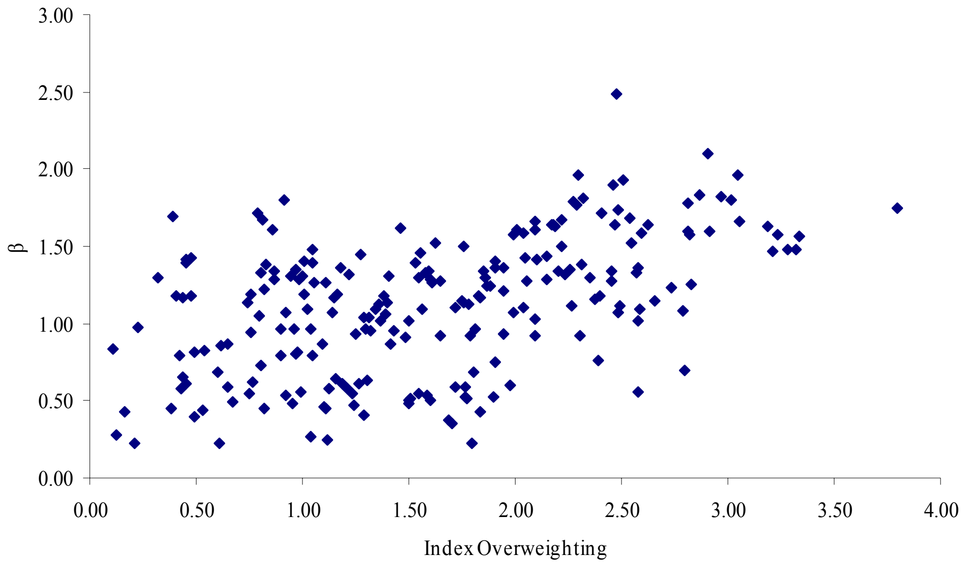


Figure 4

Comovement and Nikkei 225 index overweighting: full cross-section of index stocks

Comovement of stock returns with returns of other stocks in the Nikkei 225 are plotted against overweighting of the stock in the index. Comovement is measured as the slope (β_{EWN225}) of a time series regression of stock returns on the equal weighted return of stocks remaining in the Nikkei 225 throughout the sample period, or as the R^2 from this regression. Index overweighting is defined as the log of one plus the ratio of a stock's weight in the Nikkei 225 to the weight of the stock in the value weighted TOPIX index. Index overweighting is measured one day before the start of the sample of returns used to estimate comovement. Panel A plots comovement measure β_{EWN225} against index overweighting. Panel B plots comovement measure R^2 against index overweighting. Both plots show comovement for the entire sample of Nikkei 225 stocks on September 1, 1998.

Panel A. β_{EWN225} (N=225)



Panel B. R^2 (N=225)

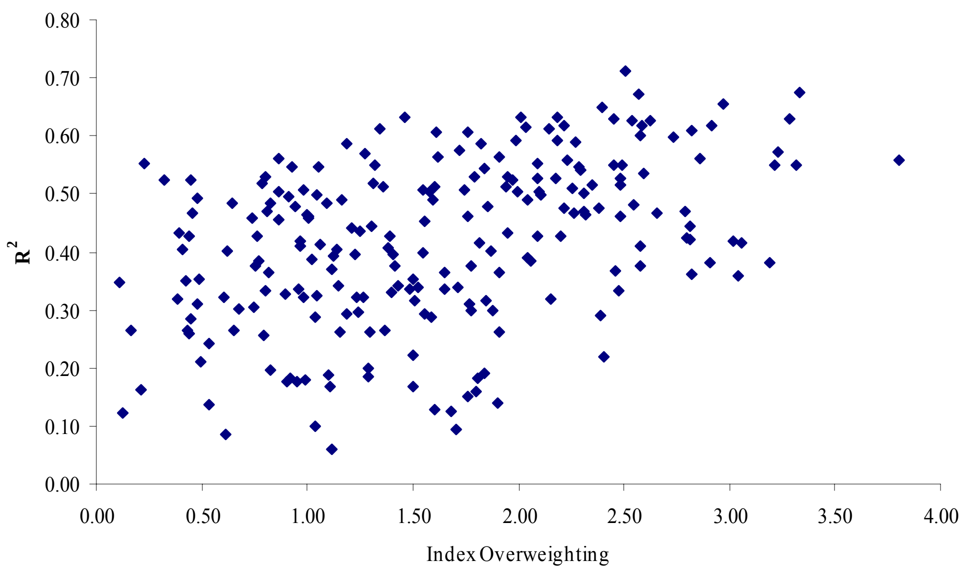


Figure 5

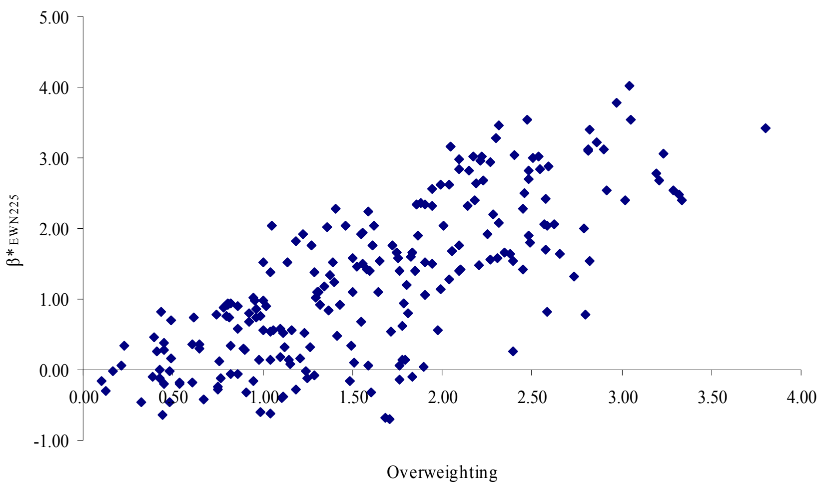
Comovement and Nikkei 225 index overweighting

Measures of comovement of stock returns with returns of other stocks in the Nikkei 225 and with returns of stocks outside of the Nikkei 225 are plotted against overweighting of the stock in the index. The figure plots the slope coefficients β_{EWN225} and β_{EWN225} from the regression

$$R_{it} = \alpha_i + \beta^*_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it}$$

against index overweighting, measured one day before the start of the sample of returns used to estimate comovement. Index overweighting is defined as the log of one plus the ratio of a stock's weight in the Nikkei 225 to the weight of the stock in the value weighted TOPIX index. Regression slope parameters are estimated using 100 days of data. Panel A plots β_{N225} for every stock in the sample on day September 1, 1998, the midpoint of the sample, against the index overweighting 100 days before. Panel B plots regression slope parameter β_{TOPIX} against the index overweighting 100 days before.

Panel A. Comovement with index stocks (β^*_{EWN225})



Panel B. Comovement with non-index stocks (β_{TOPIX})

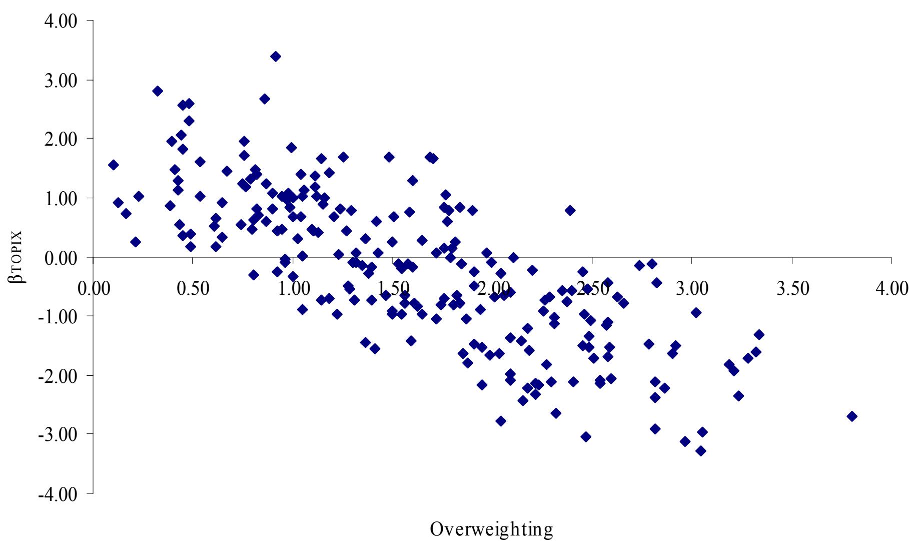


Table I

Descriptive Statistics

Cross-sectional mean, standard deviation, and extreme values of time series averages of selected variables. The main sample of stocks includes 298 stocks that were members of the Nikkei 225 index for at least one day during the period from September 1, 1993 through September 1, 2003. This period includes a total of 2609 trading days. The second sample includes 1,458 Japanese traded stocks that provided at least two years of returns data and which were not members of the index between September 1993 and September 2003. The length of the time series is the number of days for which each stock provides volume and price data. The fraction of the sample in Nikkei 225 is the percentage of the time that the stock was a member of the Nikkei 225 stock index. Daily return is the time series average of returns for each stock. Turnover is the average daily trading volume expressed as a percentage of total shares outstanding. Size is equal to the time series average of market capitalizations, in millions of yen. The price-to-book value is the share price divided by the book value per share for the appropriate financial year end, adjusted for capital changes. Leverage is the ratio of long-term debt to common (book) equity. All data are collected from Datastream. The history of index membership is constructed using the index membership changes given on the *Nihon Keizai Shimbun* webpage.

	Mean	**SD**	**Min**	**Max**
Main Sample (N=298)				
Length of time series (days)	2,451	455	109	2,609
Fraction of sample in Nikkei (%)	75.50	32.16	4.18	100.00
Daily Return (%)	0.01	0.04	-0.14	0.26
Daily turnover (%)	1.00	5.57	0.02	68.47
Size (¥ million)	910,229	1,797,490	19,759	18,884,907
P/B	2.39	3.67	0.77	55.87
Leverage	1.78	3.87	0	42.40
Other stocks (N=1,458)				
Length of time series (days)	2154	780	0	2609
Daily Return (%)	0.00	0.10	-0.81	2.20
Daily turnover (%)	0.14	0.37	0.00	12.64
Size (¥ million)	145,698	581,337	1,633	9,912,280
P/B	2.34	3.85	0	79.19
Leverage	0.92	2.42	0	34.10

Table II

Nikkei 225 index composition and weights

Correspondence between prices and Nikkei 225 weights for the 298 stocks present in the Nikkei 225 stock index for at least one day between September 1, 1993 and September 1, 2003. The value of the Nikkei 225 stock index is given by

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

where D_t is the Nikkei 225 divisor, $P_{i,t}$ is the price of stock i on day t, and F_i is the face value of stock i ranging from 50 to 50,000. The table reports average prices of stocks, where the price is taken from the first day of the sample in which a stock is present in the index. The next column reports mean stock price normalized by face value - this is the form in which prices enter the Nikkei 225 index calculation. The next column reports the average weight in the Nikkei index, the ratio of the face value adjusted price to the sum of the face value adjusted prices. The second-to-last column reports the average ratio of the Nikkei index weight to the weight in the market value weighted TOPIX index. The final column reports the average of the log of the ratio in the previous column. This is the measure of overweighting used in the paper. Reported averages are broken down by face value (50, 500, 5000, or 50000) and by additions (stocks not present in the index on September 1, 1993), deletions (stocks not present in the index on September 1, 2003), and remainders (stocks present in the index throughout the sample period).

Sample	N	Price	$\frac{\text{Price}}{\text{Face Value}/50}$	w^{N225}	$\frac{w^{N225}}{w^{VW}}$	Overweighting $\log\left(1+\frac{w^{N225}}{w^{VW}}\right)$
Full Sample	298	48,586	1,518	0.59	5.70	1.55
Face value 50	280	1,491	1,491	0.58	6.00	1.62
Face value 500	4	3,320	332	0.18	0.83	0.46
Face value 5,000	2	1,098,500	10,985	4.42	5.83	1.92
Face value 50,000	12	987,583	988	0.34	0.38	0.30
Additions	73	182,122	3,325	1.04	3.63	1.25
Face value 50	58	3,606	3,606	1.08	4.24	1.41
Face value 500	2	3,365	337	0.21	1.51	0.77
Face value 5,000	2	1,098,500	10,985	4.42	5.83	1.92
Face value 50,000	11	989,273	989	0.33	0.41	0.32
Deletions	73	14,345	1,128	0.47	10.22	1.92
Face value 50	71	1,140	1,140	0.48	10.47	1.96
Face value 500	1	4,260	426	0.31	2.79	1.33
Face value 5,000	0					
Face value 50,000	1	962,000	962	0.29	0.13	0.12
Remainders	164	6,882	944	0.45	4.34	1.47
Face value 50	161	951	951	0.45	4.42	1.49
Face value 500	2	3,275	328	0.16	0.16	0.14
Face value 5,000	0					
Face value 50,000	1	969,000	969	0.46	0.11	0.11

Table III

Nikkei 225 index overweighting and comovement for additions and deletions

Cross-sectional regressions of the change in comovement of Nikkei 225 additions and deletions on the overweighting of the stock in the Nikkei 225 stock index

$$\Delta Comove = a + b \cdot OW_{t-1} + u$$

ΔComove denotes the change in comovement, alternately defined as the change in the beta from a univariate regression of stock return on the equal weighted return of the Nikkei 225 (Δβ Univariate), the change in the R^2 from this regression (ΔR^2 Univariate), the change in the coefficient on the equal weighted return from a multivariate regression of stock return on the Nikkei equal weighted return and the return on the value weighted TOPIX index ($\Delta\beta_{EWN225}$ Multivariate), and the change in the coefficient on the TOPIX return from this multivariate regression ($\Delta\beta_{TOPIX}$ Multivariate). *OW* denotes the overweighting in the Nikkei 225 index, defined as the log of one plus the ratio of index weight to the weight the stock would have taken in the value weighted TOPIX index. For the additions, overweighting is based on closing prices on the day after the stock is added to the index. For the deletions, overweighting is defined as the closing price of the stock on the day before it is removed from the index. Results are shown separately for the 61 additions, 39 deletions, and for the combined sample of additions and deletions. In the combined sample regressions on the right-hand-side of the Table, the sign on the change in comovement is inverted for the deletions.

		Additions (N=61)			Deletions (N=39)			Combined Sample (N=100)		
ΔComove:	Dependent variable measures change in comement with:	b	[t-stat]	R^2	b	[t-stat]	R^2	b	[t-stat]	R^2
$\Delta\beta$ (Univariate)	Index stocks	0.48	[5.25]	0.32	-0.18	[-1.54]	0.06	0.27	[4.85]	0.19
ΔR^2 (Univariate)	Index stocks	0.27	[3.44]	0.17	-0.27	[-2.66]	0.16	0.69	[10.29]	0.52
$\Delta\beta_{EWN225}$(Multivariate)	Index stocks	0.47	[4.41]	0.25	-0.31	[-2.03]	0.10	0.39	[5.89]	0.26
$\Delta\beta_{TOPIX}$ (Multivariate)	Non-index stocks	-0.17	[-1.35]	0.03	0.45	[4.19]	0.32	-0.30	[-4.53]	0.17

Table IV

Nikkei 225 overweighting and comovement

Average parameter estimates and Fama Macbeth (1973) t-statistics from rolling cross sectional regressions of comovement on index overweighting at the start of the period.

$$\beta_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it}$$

$$R^2_{it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it}$$

β denotes the slope parameter from a regression of stock returns on the equal weighted return of stocks that were in the Nikkei 225 index for the entire sample period, and R^2 denotes the R^2 from this univariate regression. First stage regression estimates for period *t* are computed using returns from the interval [*t-99,t*] and are estimated for each stock for each 100 day interval. The table shows average results from second stage cross-sectional regressions of β and R^2 on lagged independent variables. These include Nikkei 225 overweighting at the start of the period (*OW*), defined as the log of one plus the ratio of index weight to the weight the stock would have taken in a value weighted index; the price-to-book ratio (*P/B*); and the log of market value (*Size*). All independent variables are measured at *t-100*, with the exception of the price-to-book ratio, which is measured in December of the previous year. In the bottom four lines of each panel, each cross-section is limited to include only the stocks that were in the Nikkei 225 index at the start and end of the period. Overweighting is defined to be zero for firms that were not in the index at the start of the period. The table also indicates the average number of firms in each cross section (N), as well as the time series average of R^2 from the cross sectional regressions. Panel A shows these results for daily measures of comovement, estimated on 26 non-overlapping intervals of 100 days. Panel B shows these results estimated using 10 non-overlapping windows of 50 weeks.

Y	Sample	*a*	[t-stat]	*b*	[t-stat]	*c*	[t-stat]	*d*	[t-stat]	N	R^2
Panel A: Daily comovement of stock returns											
β_{EWN225}	Full sample	0.723	[23.96]	0.215	[10.14]					275	0.272
β_{EWN225}	Full sample	1.553	[10.24]	0.161	[10.13]	-0.007	[-1.53]	-0.060	[-5.15]	275	0.322
R^2	Full sample	0.230	[13.89]	0.084	[12.45]					275	0.280
R^2	Full sample	0.049	[0.89]	0.094	[14.52]	-0.002	[-1.51]	0.014	[3.02]	275	0.322
β_{EWN225}	Index members	0.719	[24.38]	0.218	[10.17]					222	0.213
β_{EWN225}	Index members	2.451	[10.19]	0.047	[1.83]	-0.003	[-0.82]	-0.116	[-6.97]	222	0.299
R^2	Index members	0.272	[13.32]	0.061	[7.43]					222	0.146
R^2	Index members	0.348	[4.68]	0.053	[6.39]	-0.001	[-1.19]	-0.005	[-0.94]	222	0.178
Panel B: Weekly comovement of stock returns											
β_{EWN225}	Full sample	0.778	[11.79]	0.175	[3.85]					268	0.180
β_{EWN225}	Full sample	1.502	[8.17]	0.141	[6.34]	0.001	[3.05]	-0.053	[-3.74]	268	0.220
R^2	Full sample	0.283	[9.44]	0.083	[5.37]					268	0.204
R^2	Full sample	0.327	[4.63]	0.083	[10.40]	0.001	[3.26]	-0.004	[-0.66]	268	0.232
β_{EWN225}	Index members	0.773	[11.58]	0.180	[3.98]					218	0.148
β_{EWN225}	Index members	3.492	[12.67]	-0.068	[-3.92]	0.000	[-0.31]	-0.182	[-9.03]	218	0.224
R^2	Index members	0.317	[8.73]	0.067	[4.11]					218	0.126
R^2	Index members	1.246	[10.69]	-0.019	[-3.90]	0.001	[2.01]	-0.062	[-7.49]	218	0.189

Table V

Nikkei 225 overweighting and conditional comovement

In the first stage, I jointly estimate the conditional comovement of a stock's return with the equal weighted return of the stocks present in the Nikkei 225 index for the entire sample period, and the conditional comovement of the return with the return on the TOPIX value weighted index.

$$R_{it} = \alpha_i + \beta^*_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it}$$

The table reports average parameter estimates and Fama Macbeth (1973) t-statistics from second stage rolling cross sectional regressions of conditional comovement on index overweighting at the start of the period and a set of control variables

$$\beta^*_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it}$$

$$\beta_{TOPIX,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot (P/B)_{it-1} + d_t \cdot Size_{it-1} + u_{it}$$

where β_{EWN225} and β_{TOPIX} are the conditional comovement estimates from the first stage regressions. The independent variables in the second stage regressions include Nikkei 225 overweighting at the start of the period (*X*), defined as the log of one plus the ratio of index weight to the weight the stock would have taken in a value weighted index; the price-to-book ratio (*P/B*); and the log of market value (*Size*). All independent variables are measured at *t-100*, except for the price-to-book ratio, measured in December of the previous year. For each cross-section, the sample includes all firms present in the Nikkei index at the beginning and end of the estimation period (noted by Sample = Index Members), or all firms in the sample whose index membership did not change during the estimation period (noted by Sample = Full Sample). The table indicates the average number of firms in each cross section (N), as well as the time series average of R^2 from the cross sectional regressions.

Y	Sample	*a*	[t-stat]	*b*	[t-stat]	*c*	[t-stat]	*d*	[t-stat]	N	R^2
Panel A: Daily conditional comovement with Nikkei 225 stocks											
β^*_{EWN225}	Full sample	-0.035	[-0.48]	0.727	[13.71]					275	0.447
β^*_{EWN225}	Full sample	3.799	[16.80]	0.523	[13.34]	-0.014	[-2.59]	-0.281	[-17.25]	275	0.597
β^*_{EWN225}	Index members	-0.159	[-2.30]	0.800	[16.28]					222	0.390
β^*_{EWN225}	Index members	6.722	[17.28]	0.140	[2.27]	-0.013	[-2.56]	-0.464	[-17.72]	222	0.612
Panel B: Daily conditional comovement with stocks outside of Nikkei 225											
β_{TOPIX}	Full sample	1.028	[12.64]	-0.695	[-11.96]					275	0.358
β_{TOPIX}	Full sample	-2.977	[-12.38]	-0.483	[-11.83]	0.010	[1.36]	0.294	[15.88]	275	0.503
β_{TOPIX}	Index members	1.178	[13.53]	-0.780	[-13.06]					222	0.330
β_{TOPIX}	Index members	-5.598	[-12.22]	-0.133	[-1.87]	0.012	[1.67]	0.457	[14.85]	222	0.528

Table VI

Nikkei 225 index overweighting and comovement: Expanded sample

Average parameter estimates and Fama Macbeth (1973) t-statistics from rolling cross sectional regressions of comovement on index overweighting at the start of the period

$$\beta_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + u_{it}$$

$$R^2_{it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + u_{it}$$

β denotes the slope parameter from a univariate regression of stock returns on the equal weighted return of stocks that were in the Nikkei 225 index for the entire sample period, and R^2 denotes the R^2 from this univariate regression. First stage regression estimates for period *t* are computed using returns from the interval [*t-99,t*] and are estimated for each stock for each 100 day interval. The table shows average results from second stage cross-sectional regressions of β and R^2 on lagged independent variables. These include Nikkei 225 overweighting at the start of the period (*X*), defined as the log of one plus the ratio of index weight to the weight the stock would have taken in a value weighted index; the price-to-book ratio (*P/B*); and the log of market value (*Size*). All independent variables are measured at *t-100*, except for the price-to-book ratio, which is measured in December of the previous year. In the first two lines, each cross-section is limited to include only the stocks that were in the Nikkei 225 index at the start and end of the period. In the following two lines, this restriction is dropped and each cross-section includes all firms for which data was available. Overweighting is defined to be zero for firms that were not in the index at the start of the period. The table also indicates the average number of firms in each cross section (N), as well as the time series average of R^2 from the cross sectional regressions. Panel A shows these results for daily measures of comovement, estimated on 26 non-overlapping intervals of 100 days. Panel B shows these results estimated using 10 non-overlapping windows of 50 weeks.

Y	***a***	**[t-stat]**	***b***	**[t-stat]**	***c***	**[t-stat]**	***d***	**[t-stat]**	***e***	**[t-stat]**	**N**	**R^2**
Panel A: Daily comovement of stock returns												
β_{EWN225}	0.578	[27.18]	0.295	[22.06]							1416	0.232
β_{EWN225}	0.573	[26.79]	0.218	[10.17]	0.146	[5.74]					1416	0.239
β_{EWN225}	0.666	[5.97]	0.202	[8.07]	0.190	[4.17]	0.003	[3.41]	-0.010	[-0.98]	1416	0.264
Panel B: Weekly comovement of stock returns												
β_{EWN225}	0.709	[17.92]	0.217	[8.50]							1450	0.096
β_{EWN225}	0.708	[17.69]	0.179	[3.89]	0.074	[1.30]					1445	0.100
β_{EWN225}	1.421	[7.64]	0.060	[1.45]	0.366	[5.06]	0.005	[1.68]	-0.067	[-3.89]	1445	0.146

Table VII

Cross-sectional determinants of comovement with stocks outside of Nikkei 225 index: Expanded Sample

In the first stage, I jointly estimate the comovement of the stock return with the equal weighted return of the stocks in the Nikkei 225 index and with the return on the TOPIX value weighted index.

$$R_{it} = \alpha_i + \beta^*_{EWN225,it} \cdot R_{EWN225,t} + \beta_{TOPIX,it} \cdot R_{TOPIX,t} + \varepsilon_{it}$$

These regressions are run on a rolling basis for each stock, using prior return data from [t-100,t-1]. The table reports the results of second stage Fama Macbeth (1973) regressions of beta n225 and beta topix, the measures of comovement, on lagged independent variables.

$$\beta^*_{EWN225,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + u_{it}$$

$$\beta_{TOPIX,it} = a_t + b_t \cdot OW_{it-1} + c_t \cdot 1_{OW_{it-1}>0} + d_t \cdot (P/B)_{it-1} + e_t \cdot Size_{it-1} + v_{it}$$

The independent variables include a constant, the lagged overweighting in the Nikkei 225 index (defined to be equal to zero for non-index stocks), the price-to-book ratio, and the log of market value. All independent variables are measured at *t-100*. The table reports average coefficients from these cross-sectional regressions, together with their associated t-statistics.

Y	*a*	[t-stat]	*b*	[t-stat]	*c*	[t-stat]	*d*	[t-stat]	*e*	[t-stat]	N	R^2
Panel A: Daily conditional comovement with Nikkei 225 stocks												
β^*_{EWN225}	0.106	[2.54]	0.664	[20.50]							1,416	0.295
β^*_{EWN225}	0.115	[2.78]	0.800	[16.28]	-0.275	[-5.06]					1,416	0.301
β^*_{EWN225}	1.135	[4.05]	0.661	[9.43]	0.122	[1.15]	-0.001	[-0.64]	-0.097	[-4.21]	1,416	0.364
Panel B: Daily conditional comovement with stocks outside of Nikkei 225												
β_{TOPIX}	0.634	[15.25]	-0.496	[-14.65]							1,416	0.141
β_{TOPIX}	0.614	[14.71]	-0.780	[-13.06]	0.564	[7.14]					1,416	0.158
β_{TOPIX}	-0.614	[-2.21]	-0.616	[-8.06]	0.091	[0.76]	0.006	[1.93]	0.115	[5.07]	1,416	0.232