NBER WORKING PAPER SERIES

COMOVEMENT REVISITED

Honghui Chen
Vijay Singal
Robert F. Whitelaw

Working Paper 21281
http://www.nber.org/papers/w21281

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2015

We would like to thank Oleg Rytchkov and seminar participants at Temple University, the University of Texas, Austin, Virginia Tech, the 2014 FMA meetings, and the 2013 SFA meetings, particularly the discussant Ferhat Akbas, for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Comovement Revisited
Honghui Chen, Vijay Singal, and Robert F. Whitelaw
NBER Working Paper No. 21281
June 2015
JEL No. G14

ABSTRACT

Recent evidence of excessive comovement among stocks following index additions (Barberis, Shleifer, and Wurgler, 2005) and stock splits (Green and Hwang, 2009) challenges traditional finance theory. Based on a simple model, we show that the bivariate regressions relied upon in the literature often provide little or no information about the economic magnitude of the phenomenon of interest, and the coefficients in these regressions are very sensitive to time-variation in the characteristics of the return processes that are unrelated to excess comovement. Instead, univariate regressions of the stock return on the returns of the group it is leaving (e.g., non-S&P stocks) and the group it is joining (e.g., S&P stocks) reveal the relevant information. When we reexamine the empirical evidence using control samples matched on past returns and compute Dimson betas, almost all evidence of excess comovement disappears. The results in the literature are consistent with changes in the fundamental factor loadings of the stocks. One key element to understanding these striking results is that, in both the examples we study, the stocks exhibit strong returns prior to the event in question. We document the heretofore unknown empirical regularity that winner stocks exhibit increases in betas. Thus, much of the apparent excess comovement is just a manifestation of momentum.

Honghui Chen
University of Central Florida
College of Business Administration
4000 Central Florida Blvd
Orlando, FL 32816
honghui.chen@bus.ucf.edu

Vijay Singal
Virginia Tech
Pamplin College of Business
Blacksburg, VA 24061
vs@vt.edu

Robert F. Whitelaw
New York University
Stern School of Business
44 West 4th Street, Suite 9-190
New York, NY 10012-1126
and NBER
rwhitela@stern.nyu.edu

1. Introduction

In a perfect and frictionless financial market, asset prices change to reflect new information about future cash flows and discount rates. To the extent that there are common factors affecting either cash flows or discount rates, asset prices will move together to reflect innovations in such common factors.

However, there is growing evidence that prices move together for reasons that are seemingly unrelated to fundamentals. Evidence of this excess comovement has been found among S&P500 index additions and deletions (Vijh, 1994; Barberis, Shleifer, and Wurgler, 2005), changes in S&P500 value and growth indices (Boyer, 2011), changes in the Nikkei 225 index (Greenwood and Sosner, 2007), changes in UK indices (Mase, 2008), changes in Nikkei 225 index weights (Greenwood, 2008), additions to many national market indices (Claessens and Yafeh, 2011), stock splits (Green and Hwang, 2009), stocks with correlated trading among retail investors (Kumar and Lee, 2006), stocks with corporate headquarters in the same geographic area (Pirinsky and Wang, 2006), stocks with similar institutional ownership (Pindyck and Rotemberg, 1993), stocks in closed-end country funds (Hardouvelis et al., 1994; Bodurtha et al., 1995), stocks in closed-end domestic funds (Lee et al., 1991), sovereign bonds (Rigobon, 2002), and commodity futures (Tang and Xiong, 2012).

Though excessive comovement in stock returns is attributed to several non-fundamental factors,[1] the primary explanation is an asset class effect, which is created by correlated demand unrelated to fundamentals for assets in a particular class. Theoretical models developed by Basak and Pavlova (2013), DeMarzo, Kaniel and Kremer (2004), and Barberis and Shleifer (2003), among

[1] Barberis, Shleifer, and Wurgler (2005) propose three sources of friction and investor sentiment. Excess investor demand for a particular group of securities may arise because of investor awareness (habitat) or because those stocks form an asset class that is easy to follow (category). Third, the speed of information diffusion may increase for stocks included in the index. Similar arguments are in Hou and Moskowitz (2005) and Pindyck and Rotemberg (1993). Improvement in price discovery would cause the added stock to comove more strongly with index stocks than with non-index stocks. Since it is difficult to empirically distinguish between the first two views, Greenwood (2008) combines them into a single demand-based theory, or an asset class effect. The last source of friction, quicker adjustment in prices to new information is a desirable outcome of index additions because it makes prices more efficient even though it may increase comovement. In other words, there was too little comovement in the absence of efficient information diffusion, which has now been increased to an appropriate level (Claessens and Yafeh, 2011). Other explanations relate to transactions costs at an index level versus an individual stock level. However, we focus on the asset class effect as the generally accepted source of comovement.

others, are consistent with such an asset class effect. However, the sources of this correlated demand are varied: investor behavior that causes investors to choose stocks based on styles or categories (Barberis and Shleifer, 2003); agents who care about relative wealth choosing assets held by other members of the community (DeMarzo, Kaniel and Kremer, 2004); or institutional investors who care about their performance relative to an index tilting their portfolios towards stocks that are in that index (Basak and Pavlova, 2013).

Two papers, von Drathen (2013) and Kasch and Sarkar (2013), challenge the empirical evidence mentioned above in the context of two specific events, FTSE 100 and S&P500 index turnover, respectively. They both point out that these events coincide with changes in fundamentals. Our focus is on providing a more general view of the issue and regression results in the existing literature and on understanding the mechanisms that underlie the link between momentum and comovement, as explained below.

In this paper, we reexamine the evidence on comovement, focusing on two studies that document what appears to be strong support for this phenomenon, but in apparently unrelated contexts. The first is Barberis, Shleifer, and Wurgler (2005), which is considered a classic paper on comovement. Their sample consists of stocks that enter or leave the S&P500, an event that has been used by many other studies because index changes are generally believed to have little fundamental effect on the firm being added to or deleted from the index (Chen et al., 2004; Elliott et al., 2006). Their hypothesis is that stocks in the index comove more with index stocks, whereas those not in the index comove more with non-index stocks. The second paper is Green and Hwang (2009), who study comovement before and after stock splits. Specifically, their argument is that stocks with similar price levels comove more than would be justified by fundamentals, i.e., that a stock moves more with high-priced stocks prior to a split and more with low-priced stocks after a split. As with index changes, splits appear to be useful events to study because they do not affect splitting firms in any fundamental way, although the announcement may signal private information.

In both cases, the primary evidence is in the form of differences between the coefficients in two regressions conducted before and after the event: (1) a univariate regression of the stock return on the return of the group it is joining, and (2) a bivariate regression of the stock return on the returns of both the old group and the new group. The bivariate regression results in Barberis, Shleifer, and Wurgler (2005) show that for additions to the S&P500 index, their coefficient on S&P500 returns increases dramatically after they join the index while the coefficient on non-index stocks declines. In a similar vein, the bivariate regression results in Green and Hwang (2009) show that stocks after a split load more heavily on low-priced stocks (the new group) and less on high-priced stocks (the old group).

In order to better understand the implications of the excess comovement hypothesis for stock returns, we first develop a model closely related to that of Barberis, Shleifer, and Wurgler (2005). Some implications of our model are similar to those derived in their paper, but we highlight four additional important implications.

First, the model suggests that a univariate regression of the stock return on the return of the old group after the event can be very informative, a specification not examined in Barberis, Shleifer, and Wurgler (2005) or Green and Hwang (2009).

Second, the model indicates that the results of the bivariate regressions estimated by Barberis, Shleifer, and Wurgler (2005) and Green and Hwang (2009) are extremely sensitive to small changes in parameters. The sensitivity of these types of regression coefficients has been documented in the literature (Spanos and McGuirk, 2002). Most critically for our analysis, this sensitivity implies that the interpretation of these coefficient estimates is not straightforward, and that they may well provide little or no information about the question of economic interest—how much, if at all, is excess comovement responsible for the variation in stock returns.

Third, the model shows that changes in the parameters around the events, in particular shifts in loadings on the fundamental factor, can affect the univariate regression results. For example, an increase in the beta of a stock in the sample will generate an increase in the coefficient of the stock on the new group return after the event. In other words, these empirical results are also consistent with a change in fundamental comovement not just excess comovement. Of course, this phenomenon also has implications for the univariate regression of the stock return on the old group return discussed above, and, in fact, it is this regression that allows us to distinguish between the two competing explanations.

Finally, the model shows that shifts around the event in the fundamental loadings and idiosyncratic risk of the group returns can cause significant shifts in the bivariate regression coefficients, even in a world with no excess comovement. For example, if the idiosyncratic risk of the return on one group increases, the bivariate regression will shift weight from the return of this group to that of the other. In this regression both groups serve as proxies for the fundamental factor. The magnitude of idiosyncratic risk relative to fundamental risk determines a group's quality as a proxy and thus also the relative magnitude of its coefficient.

We begin our empirical analysis by reexamining comovement following index changes. We expand the Barberis, Shleifer, and Wurgler (2005) sample period of 1976-2000 to 1976-2012, using daily data, where they report their strongest results.[2] In general, based on the two univariate regressions, we find that stocks added to the S&P500 index move more with the S&P500 index but they also move more with the old group of non-S&P index stocks. The difference in beta changes is not significant for the 1976-87 period, nor is it significant for the 2001-12 period. The difference in beta changes is, however, significant for the 1988-2000 period. As in Barberis, Shleifer, and Wurgler (2005), the bivariate regression results show a significant increase in beta relative to the S&P500 index and a significant decrease in beta relative to the old group

For the stock split sample, evidence in support of comovement is essentially non-existent when the univariate regressions are analyzed: the increase in beta between returns on splitting stocks and returns on the new group (i.e., low-priced stocks) is almost equal to the increase in beta between returns on splitting stocks and returns on the old group (i.e., high-priced stocks). The bivariate regressions again show an increase in the beta with the new group, though there is no statistically significant decrease in beta relative to the old group.

These initial empirical results for the univariate regressions indicate that it may be increases in the fundamental betas of the stocks around the events that are driving much of the results reported in the literature as excess comovement. The natural question is why do these

[2] We end the S&P additions sample one year prior to the end of our data because we need one year of data after the event to compute regression coefficients.

betas increase, i.e., what do stocks added to the S&P500 and those undergoing splits have in common? The answer is that both groups of stocks exhibit exceptional performance prior to the event. In the language of the literature on cross-sectional momentum effects, they are winners. Following the usual momentum methodology, we find that betas of winner stocks increase during the formation period and continue to increase during the holding period, before declining at longer horizons. Therefore, it is likely that at least some of the results reported by Barberis, Shleifer, and Wurgler (2005) and Green and Hwang (2009) are caused by the inclusion of momentum stocks in their samples.

For the bivariate regression results, shifts around the event in the fundamental loadings and idiosyncratic risk of the group returns can cause exactly these types of effects, even in a world with no excess comovement.

Given the apparent importance of fundamental stock betas and shifts in the characteristics of the group returns, we next turn to a more refined analysis that attempts to better measure and control for these changes. First, we improve the estimation of the betas by employing a Dimson (1979) approach to adjust for non-synchronous trading using leads and lags of the relevant indices in the regressions. Though the S&P500 index consists of some of the largest stocks in the U.S. economy, index changes are concentrated mainly among the smaller stocks in the index. Similarly, the trading frequency of stocks that split may differ from that of the stocks in either the low- or high-priced indices that we construct. We add two leads and lags of the index returns to pick up these effects.

Second, we control for the additional effects of changes in the idiosyncratic risk and fundamental factor loading of group returns on measured comovement using a matched sample approach. For each index change and stock split, we choose a firm in the same size decile that comes closest based on momentum, i.e., has a similar return over the past year. If beta changes are driven primarily by momentum, these stocks will exhibit similar changes to those in the S&P addition and stock split samples. We then adopt a difference in difference in difference approach, examining the differences in the changes of the betas before and after the event across the stocks in the original sample and the matched sample. If changes in the properties of group returns are driving the bivariate regression results, then matched stocks will exhibit similar patterns in their regression coefficients, even though they did not change groups around the event.

The empirical results from this refined analysis are striking. For both S&P500 index additions and stock splits, the original sample and matched sample stocks exhibit differences in beta changes that are not significantly different. In other words, the differences between the changes across the two univariate regressions are statistically indistinguishable for the sample and control stocks. This result is compelling evidence that the apparent excess comovement is actually driven by changes in loadings on the fundamental component of returns, not by asset class effects. The control stocks also show similar changes in bivariate regression coefficients before and after the event to which the sample stocks are subject. Thus, the properties of group returns, not excess comovement, are clearly responsible for the anomalous results in the original samples. Moreover, this result is not simply an artifact of limited statistical power. The point estimates indicate that excess comovement is not economically significant either.

A breakdown of our two adjustments, i.e., the Dimson adjustment and the matched control adjustment, shows that their importance differs dramatically for the two samples. For the stock split sample, the Dimson adjustment does little, but the momentum control is critical because these stocks exhibit very strong past performance and resulting beta changes. In contrast, for the S&P500 index addition sample, the momentum effect is somewhat weaker and both adjustments are necessary. The differential momentum effect is consistent with a significantly greater proportion of winner stocks that split than the proportion of winner stocks that are added to the S&P500 index. The Dimson adjustment becomes more important for S&P500 additions because the added stocks are among the smallest firms in the index, which can induce spurious cross-serial correlation between additions and the index, unlike for stock splits where relative sizes of splitting stocks and other stocks are not likely to be different.

The paper is organized as follows. In the next section, we introduce the model and examine its implications for univariate and bivariate regressions. Section 3 describes the data and methodology for momentum, index changes, and stock splits. Section 4 contains the main empirical results for the original sample. In Section 5, we revisit the model in light of these initial results, investigating specifically the effects of shifts in the parameters. In Section 6 we examine the link between momentum and beta changes and then reexamine the data in the light of this evidence. We perform several robustness checks in Section 7, and conclude in Section 8.

2. A Model

In order to understand the implications of the regression results reported in the literature for the economic importance of the excess comovement phenomenon, it is useful to write down a relatively simple and stylized model in which the coefficients in these regressions can be calculated in closed form. Our goal is not to fully capture reality, but rather, in the spirit of the model in Barberis, Shleifer, and Wurgler (2005), to generate some general insights and predictions that we can use to interpret the subsequent empirical results. Our model is not identical to that in Barberis, Shleifer, and Wurgler (2005), although the key predictions are similar, because we want to construct the simplest possible model that both highlights the features of the univariate and bivariate regressions that we believe are important and captures the essence of the excess comovement hypothesis.

2.1 Setup and Assumptions

Denote as y_t the return on a stock that is changing membership between groups 1 and 2 with returns x_{1t} and x_{2t}, respectively, e.g., non-S&P and S&P stocks or high-priced and low-priced stocks:

$$\begin{aligned}
y_t &= b_{yt} f_t + c_{1t} u_{1t} + c_{2t} u_{2t} + e_{yt} \\
x_{1t} &= b_{1t} f_t + u_{1t} + e_{1t} \\
x_{2t} &= b_{2t} f_t + u_{2t} + e_{2t} \\
\mathrm{var}(e_{it}) &\equiv \sigma_{eit}^2 \quad \mathrm{var}(u_{it}) \equiv \sigma_{uit}^2 \quad \mathrm{var}(f_t) \equiv \sigma_{ft}^2
\end{aligned} \tag{1}$$

where f is the fundamental, common return shock, which could easily be extended to a multi-factor context; u_i are group-specific, non-fundamental return shocks; and e_i are idiosyncratic fundamental return shocks.

For identification purposes assume

$$
\begin{aligned}
&\operatorname{cov}(u_{1t}, u_{2t}) = 0 \\
&\operatorname{cov}(u_{it}, f_t) = \operatorname{cov}(e_{jt}, f_t) = 0 \quad \forall i, j \\
&\operatorname{cov}(u_{it}, e_{jt}) = 0 \quad \forall i, j
\end{aligned}
\tag{2}
$$

That is, non-fundamental, group-specific shocks are assumed to be uncorrelated across groups; the common fundamental factor is uncorrelated with the other shocks; and the idiosyncratic, fundamental shocks are uncorrelated with the non-fundamental shocks.

The economic content of the excess comovement hypothesis is a statement about the loadings of stock *y* on the two non-fundamental, group-specific shocks, u_1 and u_2. Specifically, using underbars and overbars to denote values prior to and after the stock switches from group 1 to group 2, the theoretical predictions of this hypothesis are

$$
\begin{aligned}
&\underline{c}_{1t} = \underline{c}_1 > 0 \quad \underline{c}_{2t} = 0 \\
&\overline{c}_{1t} = 0 \quad \overline{c}_{2t} = \overline{c}_2 > 0
\end{aligned}
\tag{3}
$$

i.e., there is a zero loading on the group-specific shock of the group to which the stock does not belong, and a positive loading on the group-specific shock of the group to which the stock does belong. We also assume that all the other parameters of the model are constant in each sub-period, i.e., the periods before and after the move of stock *y* between the groups, but that they can vary across the sub-periods. As above, we use underbars and overbars to designate these parameters.

2.2 Assessing the Economic Magnitude of Excess Comovement

The goal of our empirical analysis is to assess the economic magnitude of excess comovement. In the context of the model above, a natural measure of this quantity is the fraction of the variation in stock *y*'s return that is due to excess comovement, both prior to and after the event:

$$
\frac{\underline{c}_1^2 \underline{\sigma}_{u1}^2}{\underline{\sigma}_y^2} \quad \text{and} \quad \frac{\overline{c}_2^2 \overline{\sigma}_{u2}^2}{\overline{\sigma}_y^2}
\tag{4}
$$

This measure is equivalent to the R-squared one would get if one regressed the stock return on the non-fundamental component of the corresponding group return. The analogous quantities for the group returns are

$$\frac{\underline{\sigma}_{u1}^2}{\underline{\sigma}_{x1}^2} \quad \frac{\overline{\sigma}_{u1}^2}{\overline{\sigma}_{x1}^2} \quad \frac{\underline{\sigma}_{u2}^2}{\underline{\sigma}_{x2}^2} \quad \text{and} \quad \frac{\overline{\sigma}_{u2}^2}{\overline{\sigma}_{x2}^2} \tag{5}$$

i.e., the fraction of the variance of group returns explained by the non-fundamental component

In the literature, the focus is on two regressions run both before and after the stock switches groups—a univariate regression of the stock return on the return of the group that it is joining and a bivariate regression on the returns of both groups. As we argue below, a third regression—a univariate regression of the stock return on the group that it is leaving—is also informative. Therefore, consider the following three regressions run pre- and post-switch:

$$\begin{aligned} y_t &= \alpha + \beta_1 x_{1t} + \varepsilon_t \\ y_t &= \alpha + \beta_2 x_{2t} + \varepsilon_t \\ y_t &= \alpha + \beta_{1b} x_{1t} + \beta_{2b} x_{2t} + \varepsilon_t \end{aligned} \tag{6}$$

The probability limits of the univariate regression coefficients under the model above are

$$\begin{aligned} &\underline{\beta}_1 = \frac{\underline{b}_y \underline{b}_1 \underline{\sigma}_f^2 + \underline{c}_1 \underline{\sigma}_{u1}^2}{\underline{\sigma}_{x1}^2} \quad \overline{\beta}_1 = \frac{\overline{b}_y \overline{b}_1 \overline{\sigma}_f^2}{\overline{\sigma}_{x1}^2} \\ &\underline{\sigma}_{x1}^2 = \underline{b}_1^2 \underline{\sigma}_f^2 + \underline{\sigma}_{u1}^2 + \underline{\sigma}_{e1}^2 \quad \overline{\sigma}_{x1}^2 = \overline{b}_1^2 \overline{\sigma}_f^2 + \overline{\sigma}_{u1}^2 + \overline{\sigma}_{e1}^2 \\ &\underline{\beta}_2 = \frac{\underline{b}_y \underline{b}_2 \underline{\sigma}_f^2}{\underline{\sigma}_{x2}^2} \quad \overline{\beta}_2 = \frac{\overline{b}_y \overline{b}_2 \overline{\sigma}_f^2 + \overline{c}_2 \overline{\sigma}_{u2}^2}{\overline{\sigma}_{x2}^2} \\ &\underline{\sigma}_{x2}^2 = \underline{b}_2^2 \underline{\sigma}_f^2 + \underline{\sigma}_{u2}^2 + \underline{\sigma}_{e2}^2 \quad \overline{\sigma}_{x2}^2 = \overline{b}_2^2 \overline{\sigma}_f^2 + \overline{\sigma}_{u2}^2 + \overline{\sigma}_{e2}^2 \end{aligned} \tag{7}$$

For the bivariate regression

$$
\begin{aligned}
&\underline{\beta}_{1b} = \frac{1}{1-\underline{\rho}^2_{x1,x2}}\left[\underline{\beta}_1 - \underline{\rho}_{x1,x2}\frac{\underline{\sigma}_{x2}}{\underline{\sigma}_{x1}}\underline{\beta}_2\right] \quad \underline{\beta}_{2b} = \frac{1}{1-\underline{\rho}^2_{x1,x2}}\left[\underline{\beta}_2 - \underline{\rho}_{x1,x2}\frac{\underline{\sigma}_{x1}}{\underline{\sigma}_{x2}}\underline{\beta}_1\right] \\
&\overline{\beta}_{1b} = \frac{1}{1-\overline{\rho}^2_{x1,x2}}\left[\overline{\beta}_1 - \overline{\rho}_{x1,x2}\frac{\overline{\sigma}_{x2}}{\overline{\sigma}_{x1}}\overline{\beta}_2\right] \quad \overline{\beta}_{2b} = \frac{1}{1-\overline{\rho}^2_{x1,x2}}\left[\overline{\beta}_2 - \overline{\rho}_{x1,x2}\frac{\overline{\sigma}_{x1}}{\overline{\sigma}_{x2}}\overline{\beta}_1\right] \\
&\underline{\rho}_{x1,x2} = \frac{\mathrm{cov}(\underline{x}_1,\underline{x}_2)}{\underline{\sigma}_{x1}\underline{\sigma}_{x2}} \quad \mathrm{cov}(\underline{x}_1,\underline{x}_2) = \underline{b}_1\underline{b}_2\underline{\sigma}^2_f + \mathrm{cov}(\underline{e}_1,\underline{e}_2) \\
&\overline{\rho}_{x1,x2} = \frac{\mathrm{cov}(\overline{x}_1,\overline{x}_2)}{\overline{\sigma}_{x1}\overline{\sigma}_{x2}} \quad \mathrm{cov}(\overline{x}_1,\overline{x}_2) = \overline{b}_1\overline{b}_2\overline{\sigma}^2_f + \mathrm{cov}(\overline{e}_1,\overline{e}_2)
\end{aligned} \tag{8}
$$

(see the appendix for detailed derivations).

Furthermore, if the basic parameters of the model (the weights on the common factor, the variances of the non-fundamental shocks, and the variances of the fundamental shocks) are constant over time, which is the motivation behind looking at events that are apparently unconnected to fundamentals, i.e.,

$$
\begin{aligned}
&\underline{b}_i = \overline{b}_i \equiv b_i \quad \underline{\sigma}^2_{ui} = \overline{\sigma}^2_{ui} \equiv \sigma^2_{ui} > 0 \quad \underline{\sigma}^2_{ei} = \overline{\sigma}^2_{ei} \equiv \sigma^2_{ei} \quad i = 1, 2 \\
&\underline{b}_y = \overline{b}_y = b_y \quad \underline{\sigma}^2_{ey} = \overline{\sigma}^2_{ey} = \sigma^2_{ey} \quad \underline{\sigma}^2_f = \overline{\sigma}^2_f \equiv \sigma^2_f
\end{aligned} \tag{9}
$$

then

$$
\begin{aligned}
&\underline{\beta}_1 > \overline{\beta}_1 \quad \underline{\beta}_2 < \overline{\beta}_2 \\
&\underline{\beta}_{1b} > \overline{\beta}_{1b} \quad \underline{\beta}_{2b} < \overline{\beta}_{2b}
\end{aligned} \tag{10}
$$

(again, see the appendix for details). Intuitively, when the stock switches from group 1 to group 2, it begins to move with the non-fundamental shock to group 2 and ceases to move with the non-fundamental shock to group 1; therefore, its coefficient on group 1 returns decreases and its coefficient on group 2 returns increases, both in a univariate and a bivariate context.

If we further assume that (i) the groups are fundamentally well-diversified, i.e., there is no idiosyncratic fundamental shock at the group level ($\sigma^2_{e1} = \sigma^2_{e2} = 0$), (ii) stock *y* has a loading of one on the non-fundamental group shock, i.e., $\underline{c}_1 = \overline{c}_2 = 1$, and (iii) the loadings on the fundamental shocks are all equal to unity, i.e., $b_y = b_1 = b_2 = 1$, then we duplicate the more specific results contained in Prediction 2 of Barberis, Shleifer, and Wurgler (2005):[3]

$$\underline{\beta}_{1b} = 1, \underline{\beta}_{2b} = 0 \quad \overline{\beta}_{1b} = 0, \overline{\beta}_{2b} = 1 \tag{11}$$

This result is important because it illustrates a flaw in the interpretation of the bivariate regression coefficients. From an economic standpoint, we are not directly interested in these coefficients; the key parameters are the loadings of the stock return on the various factors in equation (1) and the variances of these factors, which determine the measures of excess comovement defined in equations (4) and (5) above. However, under the assumptions outlined above, the bivariate regression coefficients are completely independent of the variances of the non-fundamental component of group and stock returns as long as these quantities are strictly positive. Thus, even when the non-fundamental component of both stock *y* and group returns is economically meaningless, in the sense that it contributes essentially nothing to the variability of returns, the bivariate coefficients appear to suggest a dramatic and economically meaningful change in the comovement properties of stock returns as a stock switches groups.

Of course, this extreme invariance result does depend on the assumed factor loadings, specifically the fact that the stock *y* and the groups load equally on both the fundamental and non-fundamental factors.[4] However, in more general settings, it is still the case that the coefficients in the bivariate regression are sensitive to small changes in the parameters of the driving processes, and their magnitudes do not reflect the quantities of ***economic*** interest. The intuition is that all reasonably well-diversified stock portfolios tend to be very highly correlated. Thus, the correlation between the returns on the two groups of stocks will be close to one. This issue is the multi-collinearity in the bivariate regression that is discussed in Barberis, Shleifer, and

---

[3] See the appendix for details. This result is not identical to that in Barberis, Shleifer, and Wurgler (2005). Specifically, their result is slightly weaker: $\underline{\beta}_{1b} = 1, \underline{\beta}_{2b} = 0 \quad 0 < \overline{\beta}_{1b} < 1, 0 < \overline{\beta}_{2b} < 1, \overline{\beta}_{1b} + \overline{\beta}_{2b} = 1$

This difference is due to the fact that Barberis, Shleifer, and Wurgler (2005) assume a multi-factor structure for fundamentals, where each group loads on a common factor and its own, unique fundamental shock. Barberis, Shleifer, and Wurgler (2005) also allow for correlation across the group-specific, non-fundamental shocks.

[4] While this appears to be a strong assumption, it is essentially equivalent to saying that stock *y* is an “average” stock in both groups 1 and 2. This assumption is unlikely to be strictly true, but it may be a reasonable first approximation.

Wurgler (2005). As they rightly point out, multi-collinearity does not affect the consistency of the estimates in OLS. But, as the example above illustrates, the magnitudes of the coefficients in the bivariate regression may tell us very little, or even nothing, about what we really want to know, i.e., how much excess comovement affects returns. This concern is especially relevant if the strong assumptions above about the stability of the parameters across the two sub-periods, which are critical in deriving the results, are not valid.

Fortunately, the coefficients in the univariate regressions isolate precisely the quantities of interest. Going back to the more general assumptions about stability of the parameters across the sub-periods, but making no assumptions about the magnitudes of the factor loadings, the differences between these coefficients pre- and post-switch are (see the appendix for details):

$$\begin{aligned}\overline{\beta}_1 - \underline{\beta}_1 &= -\frac{\underline{c}_1 \sigma_{u1}^2}{\sigma_{x1}^2} \\ \overline{\beta}_2 - \underline{\beta}_2 &= \frac{\overline{c}_2 \sigma_{u2}^2}{\sigma_{x2}^2}\end{aligned} \tag{12}$$

Thus, empirical evidence that the coefficient on the return of the group to which a stock is moving (group 2) increases after the switch would appear to be strong evidence of excess comovement. The magnitudes of these differences are also informative about the economic importance of this phenomenon. Assuming the loadings on non-fundamental group shocks equal one, which will be true on average since the shock at the group level is the value-weighted average of the shocks to the stocks within the group, these quantities are the fraction of the variation of group returns explained by excess comovement. For example, an increase of 0.1 in the beta on group 2 or a similar decrease in the beta on group 1 would indicate that 10% of the variation in group returns is due to excess comovement. Multiplying this number by the ratio of group variance to stock variance will yield the corresponding R-squared for individual stocks.

Finally, one might think that the problems in the bivariate regression are due solely to the multi-collinearity problem associated with the high correlation between the group returns. This conjecture is not true, since orthogonalizing the variables is not a complete solution. Consider, for example, a trivariate regression of the stock return on the fundamental factor and the components of the two group returns that are orthogonal to this factor—the non-fundamental factor and the idiosyncratic shock. In this regression, the magnitudes of the coefficients on these orthogonal components are relatively uninformative about the economic magnitude of excess comovement, completely so when the group returns are perfectly well-diversified. These coefficients will equal the stock's loadings on the non-fundamental shocks, c_i, but they contain no information about the variance of these shocks, σ_{ui}^2, the key terms in equations (4) and (5). In the more general setting, changes in the magnitude of idiosyncratic volatility at the group level also affect these coefficients.

3. Data and Empirical Methodology

Given these preliminary theoretical results, we turn to a reexamination of the empirical evidence in the next section, preceded in this section by a brief description of the data and the empirical methodology. The CRSP stock files at the University of Chicago and Standard and Poor's are the primary sources of data. In general, we follow the methodologies in Barberis, Shleifer, and Wurgler (2005) and Green and Hwang (2009) for constructing our tests. For index changes, we follow the methodology of Barberis, Shleifer, and Wurgler (2005) except that we use only daily data because their results are weaker with weekly and monthly data. Barberis, Shleifer, and Wurgler (2005) use additions to the S&P500 from 1976 to 2000 and deletions from 1979 to 2000, whereas our initial sample extends from 1976 to 2012 for index additions.[5] However, subperiod analysis corresponds to their subperiods. Index deletions are evaluated for robustness in Section 7. Like Barberis, Shleifer, and Wurgler (2005), we estimate betas in the pre-inclusion period using 12 months of data ending the month before the announcement of the stock's addition to the S&P500 and betas in the post-inclusion period using 12 months of data starting the month after the inclusion of the stock in the S&P500.

For stock splits, we follow the methodology in Green and Hwang (2009) and the clarifications obtained directly from the authors, though some differences in methodology persist. Like Green and Hwang (2009), our sample consists of all common stocks where the stock

[5] We limit our main analysis to index additions with share codes of 10 and 11 to remain consistent with Barberis, Shleifer, and Wurgler (2005). However, the results are similar if the sample contains all index additions.

price was $10 or more before the stock split.[6] The high-price index consists of stocks whose prices are ±25% of the price of the splitting stock just prior to the split. The low price index consists of all stocks whose price is above $5 and within ±25% of the post-split price calculated based on the pre-split price and the split ratio. The Green and Hwang (2009) sample covers the period 1971-2004. We extend this sample to 2012, and, after replicating their results for their original subperiods, we use the same subperiods as in the S&P additions sample for the subsequent analysis.

For momentum, which will become an important control variable, we follow a methodology that is similar to that in Jegadeesh and Titman (2001) and form momentum portfolios using a 12-month formation period, one skip month, and 12-month holding period. More specifically, at the end of each June from 1976 through 2011, stocks with a price of at least $10 that do not fall into the bottom size decile of NYSE stocks are assigned to 10 momentum deciles based on their cumulative returns over the preceding 252 days.[7] We estimate betas for each stock based on a rolling window of 252 days from two years before formation of momentum portfolios through two years after formation, and compare beta changes for the top and bottom momentum portfolios. Thus, betas for years -2 and -1 are estimated over rolling windows ending 504 and 252 trading days before portfolio formation, respectively. Post-formation momentum portfolio betas allow for a 21-trading day skip, and are estimated over 252 days ending 273 and 525 trading days after portfolio formation. The top return decile and the bottom return decile in the formation period are identified as winner stocks and loser stocks respectively.

4. Reexamining the Empirical Evidence

The first step in our analysis is to recreate, extend, and reexamine the univariate and bivariate regressions reported in the literature for the S&P500 index addition and stock split samples, given the insights from the model in Section 2. These are the regressions specified in equation (6), and

[6] For consistency with their results, we only include stock splits identified by CRSP with a distribution code '5523'. However, inclusion of stocks splits with a CRSP distribution code '5533' produces similar results.

[7] The sample ends in 2011 because we are evaluating beta changes up to two years after formation of momentum portfolios.

they are estimated twice, once before the event and once after. Note that the first regression, the return on the stock on the return of the group that it is leaving, is not examined in the literature. The implications of the coefficients in these regressions for the excess comovement hypothesis are discussed in Section 2.2.

The results are presented in Tables 1 and 2 for S&P500 index additions and stock splits, respectively. In each case, Panel A shows the univariate regression results and Panel B the bivariate regression results. In Panel A, the set of 3 columns beginning with the third column contain the betas relative to the old group portfolio (non-index stocks or high-priced stocks) before and after the event and the associated changes, the next set of 3 columns contain the analogous numbers relative to the new group portfolio, and the final column shows the difference between the changes in the two coefficients across the event. Panel B is organized in the same way except that the coefficients are those on the two group returns in the bivariate regressions before and after the event.

Turning first to the S&P500 additions sample, the results from the univariate regressions on the S&P500 index (the new group, i.e., group 2) for two sub-periods, 1976-87 and 1988-2000, are consistent with those reported by Barberis, Shleifer, and Wurgler (2005) in their Panel A of Table 1.[8] For 1976-87, we report a change in beta of 0.062 ($\Delta\beta_2$) based on a sample of 197 index additions compared with 0.067 in Barberis, Shleifer, and Wurgler (2005) based on a sample of 196 index additions. For 1988-2000, we and Barberis, Shleifer, and Wurgler (2005) both find an increase in beta of 0.214 after stocks are added to the S&P500 index. This increase in the difference is consistent with the excess comovement hypothesis since the latter period coincides with an increase in indexing. Interestingly, however, this difference is less than a third as large (0.071 vs. 0.214) for the very last sub-period, 2001-2012, which was not covered in the original sample, when indexing gained even more importance. Notwithstanding this anomaly, on their own, these results would naturally be interpreted, in the context of the model in Section 2, as evidence of excess comovement: The stock begins to load more heavily on the index return after it joins the index. Moreover, the economic magnitude of this effect, particularly in the 1988-2000

[8] Standard and Poor's did not publicly announce index changes until September 1976. Therefore, the first period begins in September 1976. However, for ease of reference, we term the period 1976-87.

sub-period, is large. Specifically, a coefficient of 0.214, assuming that we can interpret this average across stocks as the effect at the group level, implies that more than 20% of the variance of S&P500 returns is explained by excess comovement, i.e., the non-fundamental group-specific shock. Of course, individual stock returns are more variable than those of diversified portfolios, so the corresponding R-squareds at the stock level would be significantly smaller.

Looking at the univariate results with the non-index returns as the independent variable shows that this simple interpretation is not completely accurate. To be consistent with excess comovement, the change in the coefficient relative to the old group from before to after the stock joins the index ($\Delta\beta_1$) should be negative. That is, the stock should load less heavily on non-index returns when it is in the index, a change not examined in prior studies. Instead, we find that this change ($\Delta\beta_1$) is approximately equal in magnitude to the coefficient change for the other regression ($\Delta\beta_2$) for the 1976-87 and 2001-12 periods. Consequently, the measure of total excess comovement, the difference between these changes ($\Delta\beta_2$-$\Delta\beta_1$), is small and statistically insignificant for these two subperiods. Taken together, these results suggest that it may be changes in loadings on the fundamental factor that are more important, except for the 1988-2000 subperiod. In other words, it is not that stocks are moving more with S&P500 returns after they join the index, simply that they are moving more with all stocks.

The model in Section 2 implies that the bivariate results are unreliable in terms of assessing the economic magnitude of any excess comovement, but, for completeness, we present results from the bivariate regressions in Panel B. These results are similar to those reported in Barberis, Shleifer, and Wurgler (2005) for matching subperiods. Their bivariate regressions show an increase in the beta with the S&P index (new group) and a decrease in the beta with non-S&P500 stocks (old group). For example, for the full sample the average beta on the non-S&P group decreases by 0.305, while the beta on the S&P500 increases by 0.338.

Interestingly, these results are very different from those in the univariate regressions, where both coefficients increase. The bivariate regression coefficients may say little about the magnitude of excess comovement, but this discrepancy suggests that there are additional shifts in the model parameters across the events. Changes in the fundamental loadings of the group returns and in the idiosyncratic risk of these portfolios will affect the bivariate coefficients much more than their univariate counterparts, as we demonstrate in the next section.

The results for stock splits are reported in Table 2, first with the Green and Hwang (2009) sub-periods. The changes in beta relative to the new group reported in Table 2, columns 6-8, for matching sub-periods are very close to those reported by Green and Hwang (2009) in their Panel A of Table 2: we report a change of 0.196 for 1971-1990 with a sample of 2,350 splits compared to their change in beta of 0.204 with a sample of 2,302 splits for the same period. For the 1991-2004 period, the samples are marginally different: Green and Hwang (2009) report an increase of 0.255 in beta with a sample of 2,303 splits compared to 0.248 with a sample of 2,478 splits in this paper. The second sets of results use the subperiods in Table 1 for consistency in the following tables; the results are very similar, and there is little variation over time. As for index changes, the univariate regressions results are striking. The coefficient on low-priced stocks increases significantly after the split for all sub-periods and is consistent with the notion of excess comovement documented in the earlier studies.

We also examine the change in beta relative to the old, high-priced group before and after the split. From Panel A of Table 2, columns 3-5, we can see that $\Delta\beta_1$ is significantly positive for all sub-periods, which suggests that the beta of the splitting stock increases not only relative to the new group (low-priced stocks) but also relative to the old group (high-priced stocks). Turning to the difference in the change in betas, $\Delta\beta_2$-$\Delta\beta_1$, we find that these numbers are small. For two of the sub-periods they are negative. Although the differences of 0.03 and 0.01 are statistically significant in the 1988-2000 period and the full 1976-2012 sample, the economic magnitudes are very small and unimportant. Overall, the evidence is that the splitting stocks move more with both the old group and the new group to approximately the same extent. Thus, there is little or no reliable evidence of excess comovement following stock splits. The vast majority of the apparent effect is attributable to an increase in the fundamental beta of these stocks.

The unreliable bivariate regressions show an increase in comovement with the new group. For example, over the full period the beta on high-priced stocks falls by 0.026, while the beta on low-priced stocks increases by 0.239. However, as with the S&P500 additions sample, this discrepancy between the univariate and bivariate regression results may be an indication of shifts in the properties of the group returns in addition to the increase in the fundamental beta of the individual stocks suggested by the univariate regression results.

5. Model Implications and Parameter Instability

The empirical results in Section 4 suggest that the fundamental betas of the stocks in the two samples are increasing around the event. Moreover, there are more complex patterns in both the univariate and bivariate coefficients that are potentially consistent with changes in the parameters of the model that are not associated with excess comovement. Specifically, in one subperiod the S&P additions sample shows an increase in the relative beta on the S&P500 in the univariate regression, and both samples show shifts in the loadings from the group that the stock is leaving to the group that it is joining in the bivariate regressions.

In this section, we again turn to the model from Section 2 to consider in more detail the effects of three forms of parameter instability that can potentially explain these results—(1) changes in the fundamental betas of the stocks, (2) changes in the idiosyncratic risk of group returns, and (3) changes in the fundamental betas of group returns. Throughout this analysis we assume that there is no excess comovement at all, i.e.,

$$\underline{\sigma}_{ui}^2 = \overline{\sigma}_{ui}^2 = 0 \quad i = 1, 2 \tag{13}$$

so that all the changes in the coefficients are driven by changes in fundamentals.

While the univariate and bivariate are available in closed form, as shown in Section 2, it is easier to get the economic intuition for the effects of parameter instability in the context of some simple numerical examples, where the parameter values are chosen to be representative of those in the data.[9] We start with a base case and examine variants of this example in the subsections to follow. For the base case we assume (1) no parameter instability, i.e., the parameters are the same before and after the group switch, and (2) perfect symmetry across the

[9] In our stylized model there is a single unobservable fundamental factor. To calibrate this model we use the value-weighted CRSP portfolio to proxy for this factor. The properties of the group returns, i.e., their betas with respect to this factor and their residual risk, vary across the two samples and across the two groups within each sample, so for ease of exposition we use parameter values within the range spanned by the data.

two groups, i.e., the parameters governing the two group returns are the same. More specifically, we assume

$$\underline{b}_i = \overline{b}_i = 1 \quad \underline{\sigma}_{ei} = \overline{\sigma}_{ei} = 0.2\% \quad i = 1, 2$$
$$\underline{b}_y = \overline{b}_y = 1 \quad \underline{\sigma}_{ey} = \overline{\sigma}_{ey} = 1.73\% \quad \underline{\sigma}_f = \overline{\sigma}_f = 1\% \tag{14}$$

with volatilities computed on a daily basis. These daily volatilities imply an annualized volatility of the fundamental factor of 15.9% and annualized total (idiosyncratic) volatilities at the group and stock levels of 16.2% (3.2%) and 31.7% (27.5%), respectively. The qualitative nature of the results below are not affected by the precise parameterization. For convenience, we further assume that the idiosyncratic shocks at the group level are uncorrelated

$$\text{cov}(\underline{e}_1, \underline{e}_2) = \text{cov}(\overline{e}_1, \overline{e}_2) = 0 \tag{15}$$

This covariance influences the bivariate regression coefficients, but this assumption has no qualitative effect on the key results, i.e., the changes in coefficients across the event.

The resulting univariate and bivariate regression coefficients are

$$\underline{\beta}_1 = \underline{\beta}_2 = \overline{\beta}_1 = \overline{\beta}_2 = 0.962 \quad \overline{\beta}_1 - \underline{\beta}_1 = \overline{\beta}_2 - \underline{\beta}_2 = 0.000$$
$$\underline{\beta}_{1b} = \underline{\beta}_{2b} = \overline{\beta}_{1b} = \overline{\beta}_{2b} = 0.490 \quad \overline{\beta}_{1b} - \underline{\beta}_{1b} = \overline{\beta}_{2b} - \underline{\beta}_{2b} = 0.000 \tag{16}$$

These base case results and the associated parameter inputs are summarized in the first row of Table 3, Panels B and A, respectively, along with the corresponding inputs and results for the three other numerical examples discussed in Sections 5.1-5.3 in the succeeding rows. Due to the assumptions of parameter stability and symmetry, the coefficients are identical across the two groups and across the pre- and post-event period. The univariate coefficients are slightly less than 1 because idiosyncratic risk at the group level causes a slight attenuation of the coefficient. In other words, the group return is proxying for the fundamental factor, but it is not a perfect proxy because there is a small amount of idiosyncratic risk. In the bivariate regressions, the fundamental loading is split equally across the two groups with similar but somewhat smaller attenuation.

5.1 Changes in Stock Betas

First, consider the case where the loading of stock *y* on the fundamental factor, b_{yt}, is allowed to vary across the sub-periods but all the other parameters are kept at their values in equations (14) and (15). Specifically, assume

$$\underline{b}_y = 1.0 \quad \overline{b}_y = 1.2 \tag{17}$$

i.e., the fundamental loading of the stock increases by 20% after the event.

The resulting univariate and bivariate regression coefficients are[10]

$$\begin{array}{lll} \underline{\beta}_1 = \underline{\beta}_2 = 0.962 & \overline{\beta}_1 = \overline{\beta}_2 = 1.154 & \overline{\beta}_1 - \underline{\beta}_1 = \overline{\beta}_2 - \underline{\beta}_2 = 0.192 \\ \underline{\beta}_{1b} = \underline{\beta}_{2b} = 0.490 & \overline{\beta}_{1b} = \overline{\beta}_{2b} = 0.588 & \overline{\beta}_{1b} - \underline{\beta}_{1b} = \overline{\beta}_{2b} - \underline{\beta}_{2b} = 0.098 \end{array} \tag{19}$$

The increase in the fundamental loading of the stock from 1.0 to 1.2 shows up almost one for one in the regression coefficients, with this change being split equally between the two bivariate coefficients.

For the univariate regressions, these results coincide closely with those in Table 2, Panel A for the stock split sample. With the exception of the 1988-2000 sample period, they also look like those in Table 1, Panel A for the S&P additions sample. In other words, there is clear evidence of an increase in the fundamental loadings of the stocks across the events. However, the bivariate results paint a more complex picture in both cases. It is clearly not the case that this increase shows up equally in both coefficients in these regressions. Thus, for the bivariate regression results to be consistent with the absence of excess comovement, there must be other shifts in the parameters. We turn next to the effect of changes in the idiosyncratic risk of the group returns.

[10] For ease of reference, we tabulate these results in the second row of Table 3, Panels A and B.

5.2 Changes in Group Idiosyncratic Risk

Let us return to the base case parameter values, with the exception that we now allow the idiosyncratic risk of the group 1 returns to vary across the event. Specifically,

$$\underline{\sigma}_{e1}^{2} = 0.20\% \quad \overline{\sigma}_{e1} = 0.24\% \quad \underline{\sigma}_{e2}^{2} = \overline{\sigma}_{e2}^{2} = 0.20\% \tag{20}$$

i.e., the idiosyncratic volatility of group 1 returns increases by 20%. Note that because the group is well-diversified and thus idiosyncratic risk is small to begin with, this increase moves the total annualized volatility of group 1 returns from 16.2% to only 16.3%.

The resulting univariate regression coefficients are

$$\underline{\beta}_1 = 0.962 \quad \overline{\beta}_1 = 0.946 \quad \underline{\beta}_2 = \overline{\beta}_2 = 0.962 \quad \overline{\beta}_1 - \underline{\beta}_1 = -0.016 \quad \overline{\beta}_2 - \underline{\beta}_2 = 0.000 \tag{21a}$$

and the bivariate regression coefficients are

$$\underline{\beta}_{1b} = \underline{\beta}_{2b} = 0.490 \quad \overline{\beta}_{1b} = 0.400 \quad \overline{\beta}_{2b} = 0.577 \quad \overline{\beta}_{1b} - \underline{\beta}_{1b} = -0.090 \quad \overline{\beta}_{2b} - \underline{\beta}_{2b} = 0.086 \tag{21b}$$

(see the third row of Table 3, Panels A and B). The group one return is now a slightly poorer proxy for the fundamental factor after the event. This effect shows up in the univariate regression as a small decline of 0.016 in the group one beta. However, the effects on the bivariate regression coefficients are much more dramatic. After the event, the regression shifts substantial weight from the group one return to the group two return. Even though the volatility of the group one return has only gone up slightly, this return is highly correlated with the group two return, so even a small deterioration in its ability to proxy for the fundamental factor causes a large move in the coefficients. Specifically, the coefficient on the group one return declines by 0.1, more than 5 times the magnitude of the move in its univariate counterpart, and in sharp contrast to the result in Section 5.1 above where, as expected, the bivariate coefficients move by about half as much as those in the univariate regressions. There is also a roughly corresponding increase in the coefficient on the group two return. Note that we obtain these spurious results with bivariate regressions though we explicitly assumed no excess comovement in the setup.

There are two additional features to note about changes in the idiosyncratic volatility of group returns. First, at these parameter values the magnitude of the percentage change in the bivariate coefficients is approximately equal to the percentage change in idiosyncratic volatility—20% in the numerical example above. Second, a qualitatively and quantitatively similar effect arises if the idiosyncratic volatility of group two returns declines. The key point is that economically small movements in volatility can produce shifts in the coefficients in the bivariate regressions as documented for both the S&P500 and stock split samples. However, these shifts cannot explain the differences between the changes in the univariate coefficients in the 1988-2000 subsample for S&P500 additions. To resolve this anomaly, we next consider shifts in the fundamental betas of the group returns.

5.3 Changes in Group Betas

Finally, to see the effects of a change in the fundamental beta of the group returns, consider again the base case with parameter stability and symmetry across the groups, except that the beta of group 2 (the group that the stock is joining) changes across the event. Specifically,

$$\underline{b}_1 = \overline{b}_1 = 1.0 \quad \underline{b}_2 = 1.0 \quad \overline{b}_2 = 0.8 \tag{22}$$

i.e., the fundamental loading of the group 2 returns declines by 20% across the event.

The resulting univariate and bivariate regression coefficients, as also reported in the final row of Table 3, Panels A and B, are

$$\begin{aligned} &\underline{\beta}_1 = \overline{\beta}_1 = 0.962 \quad \underline{\beta}_2 = 0.962 \quad \overline{\beta}_2 = 1.176 \quad \overline{\beta}_1 - \underline{\beta}_1 = 0 \quad \overline{\beta}_2 - \underline{\beta}_2 = 0.215 \\ &\underline{\beta}_{1b} = \underline{\beta}_{2b} = 0.490 \quad \overline{\beta}_{1b} = 0.595 \quad \overline{\beta}_{2b} = 0.476 \quad \overline{\beta}_{1b} - \underline{\beta}_{1b} = 0.105 \quad \overline{\beta}_{2b} - \underline{\beta}_{2b} = -0.014 \end{aligned} \tag{22}$$

Given these parameter values, the increase in the univariate coefficient on group 2 (0.215) is approximately equal to the decrease in the fundamental beta of the group 2 returns (0.200). The primary effect is that the group two return is now less sensitive to the fundamental factor after the event and therefore the loading on this return must increase in order to explain the unchanged fundamental loading of the stock.

In the bivariate regression, this increase shows up as a smaller 0.105 increase in the coefficient on the group one return with little change in the coefficient on the group two return.

As in the univariate regression, the loadings are adjusting so that the fundamental loading of the stock is almost fully captured. However, after the event the regression favors the group one return as a proxy for the fundamental factor because, with a decreased beta but unchanged idiosyncratic volatility, the group two return has now become a relatively poorer proxy.

5.4 A Matched Sample Approach

Subsections 5.1-5.3 illustrate that parameter instability can generate effects on the univariate and bivariate regression coefficients similar to those seen in the data, even in our stylized model and, more importantly, in the complete absence of excess comovement. Of course, excess comovement can also generate movements in the coefficients. The question is whether we can distinguish between these competing explanations. We can potentially identify shifts in the parameters in the data that are consistent with the logic above, but it is important to remember that our numerical results are in the context of a stylized model. The real data generating processes are undoubtedly more complex. However, there is a different approach that will allow us to determine if the empirical results are driven by excess comovement. In particular, shifts in the properties of the group returns will show up in the regression results regardless of the identity of the stocks whose returns are used as the dependent variables. If we can find a sample of stocks that match the key features of the changes in properties of the stock returns in the two samples, i.e., the movements in their fundamental betas, then all the other effects associated with the group returns will show up in regressions using this matched sample. We pursue this exercise in Section 6.

6. Comovement Revisited

It would be a remarkable coincidence if selecting samples based on S&P500 index additions and stock splits was independently choosing stocks whose betas increase after the event. However, as it turns out, these two samples have something in common. The stocks in both samples have abnormally good performance before the event. This phenomenon is well known for stock splits—only companies whose stock price goes up split their stocks—but it is also intuitive for index additions—S&P is biased towards larger, better-performing stocks for inclusion in their flagship index, holding other criteria constant. Moreover, the goal of making the index representative of the market in terms of industry balance also leads to the inclusion of industries and firms within these industries that have performed relatively well.

To examine the extent of these effects, for each stock in the two samples we record in which momentum decile it falls. In other words, when stocks are ranked into 10 portfolios based on returns over the past year (i.e., from losers to winners), how many of our sample stocks are in each portfolio? These results, along with the mean and median returns of the sample stocks are reported in Table 4. If the decision to include a stock in the S&P500 or to split were independent of past returns, we would expect approximately 10% of the sample to fall in each decile. In contrast, both samples are tilted heavily towards winner stocks, with the effect being more pronounced for the split sample. For example, 57% of the split sample falls into the top 2 deciles, while the corresponding number for S&P500 additions is 37%. Average returns for these samples are 109.1% and 41.6%, although the medians are lower, suggesting a right-skewed distribution.

Given this evidence, the questions are (1) whether selecting on positive past performance can explain the beta increases that are consistent with the initial empirical results in Section 4, and (2) whether controlling for this effect eliminates the appearance of excess comovement. We look at the former question in Section 6.1 and the latter in Section 6.2.

6.1 Momentum and Beta

In examining changes in beta following periods of good performance, we follow the momentum methodology described in Section 3. While our focus is on winners, we report the winner and loser stock betas beginning 2 years before the holding period and continuing up to 2 years after the beginning of the holding period.[11] The results, in Table 5 and Figure 1, show that betas of winner stocks increase dramatically during the formation period and continue to increase during

[11] These tests require a long trading period potentially leading to a survivorship bias. The results, however, are virtually unaffected even when shorter periods are used.

the holding period. They stabilize thereafter for a few months and begin to decline. Specifically, we find that betas of winner stocks increase from 0.976 to 1.143 (a statistically significant change of 0.167) from Year-1 to Year 0, and from 0.964 in Year-2 to 1.143 in Year 0, a statistically and economically significant increase of 0.179. The betas continue to increase further during the holding period to 1.271 (a statistically significant change of 0.128) from Year 0 to Year+1 before declining to 1.166 in Year+2.

This pattern of consistently increasing betas for stocks with high past returns has the potential to explain the results in Section 4. The betas of the stocks in the sample increase around the event in question, and therefore they comove more with all stocks after the event, both stocks in the group they are joining and stocks in the group they are leaving.

6.2 Comovement with Momentum Matched Firms (and Dimson's betas)

For the analysis in this subsection, we make two adjustments in order to better assess the magnitude of excess comovement, if any, present in the data. First, because we are using daily data, nonsynchronous trading may limit our ability to get accurate regression coefficients. To the extent that stocks do not all trade simultaneously at the end of each day, the observed return on a stock will be potentially correlated with leads and lags of the returns on a given portfolio (Denis and Kadlec (1994)). The correct adjustment for this effect in order to uncover the true regression coefficient is to sum the coefficients in a regression which includes these leads and lags (Dimson (1979)). Nonsynchronous trading is likely more important for the stock splits sample since these stocks are smaller and less liquid on average than those added to the S&P500. However, this adjustment is likely to be more important for the S&P500 additions sample when we examine changes in coefficients pre- and post-event. The intuition is that it is changes in nonsynchronous trading across the two periods that matter for examining differences in coefficients, and while there is little evidence of major liquidity effects associated with stock splits that is not true for index additions. Throughout the analysis in this section, we use two leads and lags for all portfolio returns used as independent variables.

Second, following up on the Section 6.1 results, where we find evidence of increasing betas in momentum stocks, and the matched sample logic of Section 5.4, we also compare comovement of sample stocks with a matched sample that exhibits similar momentum characteristics. Barberis, Shleifer, and Wurgler (2005) use a sample of firms matched by size and industry, but do not control for momentum, which appears to be the critical factor due to the beta patterns associated with winner stocks. Consequently, for each addition, we select a matched firm from the same size decile that is not a member of the S&P500 index and is closest in terms of lagged 252-day return to the added firm at the time of inclusion.[12] Due to the exceptional performance of some firms in the sample, a perfect match is not possible. While the average and median returns of the matched stocks are only slightly lower than those of the original sample, for stocks in the top 10 percent of the sample, the matched stocks have returns that are significantly lower, albeit still high, in some cases.[13]

Like Barberis, Shleifer, and Wurgler (2005), Green and Hwang (2009) construct a sample matched by size and industry without controlling for momentum. The matched sample that we use in this paper for stock splits controls for both size and momentum. For each stock split, we first select a group of firms from the high-priced portfolio that fall in the same size decile. Thereafter, we choose firms that are closest to the splitting firm in terms of momentum. The matched firm is the one that comes closest in price and momentum to the sample firm within the same size decile. Given the more challenging matching criteria and the more extreme positive returns of the stock split sample, it is not surprising that the match is somewhat worse than for the S&P500 additions sample. In this case, even the median return of the matched sample is more than 7% below that of the original sample, with much larger differences for stocks with the most extreme returns. In spite of this issue, it is still worth examining the results, realizing that if the magnitudes of beta changes are correlated with the magnitudes of returns, particularly for very high returns, the matched sample will not exhibit quite the same shifts in fundamentals as the original sample.

[12] In results not presented here, requiring the matched firm to be from the same industry as the sample firm does not change the results.
[13] In the interests of brevity, these results are not tabulated in the paper.

Tables 5 and 6 present the results for the S&P500 index addition and stock split samples, respectively. In both cases, Panel A provides the univariate regression results, while those for the bivariate regression are reported in Panel B. Within each panel, we first present the results for the sample of event stocks. These results are comparable to those in Tables 1 and 2, except that we now use the Dimson adjustment to estimate the coefficients. We then provide the estimation results for the matched sample. Finally, we show the difference between the original and momentum-matched samples.

For the S&P500 index additions, the Dimson adjustment alone generally accounts for more than 50% of the effect that appears in the original analysis. For example, $\Delta\beta_2$ in the most significant sub-period (1988-2000) drops from 0.214 to 0.078. Not surprisingly, this large change is primarily due to an increase in the estimated beta prior to the addition of the stock to the index. It is prior to being included in the index that the stock is likely to be less liquid, and therefore the Dimson adjustment is also likely to be more important. Looking at the differences between the coefficient changes across regressions, $\Delta\beta_2$-$\Delta\beta_1$, only in this same sub-period is the coefficient statistically positive with a value of 0.129 and a t-statistic of 2.55. However, a similar result holds for the matched sample in 1988-2000. For these firms we get a value of 0.111 with a t-statistic of 2.23. Across all sub-periods there is no single difference above 0.020 between the original and matched samples. To put it succinctly, there is absolutely no evidence of any excess comovement once we control for the momentum effect.

That said, one might legitimately wonder why, in the 1988-2000 subperiod, both the sample and matched stocks exhibit univariate regression coefficients that vary so much across S&P500 and non-S&P500 stocks. The answer, as discussed in Section 5, is a shift in fundamental parameters over the event period. First, it is important to note that the anomalous result above is confined to the years 1999 and 2000. For the other years in the subperiod, there are no statistically significant effects. However, in these two years the effect reported in Panel A of Table 6 is much larger. The explanation is a shift in the fundamental betas of the two groups of stocks, S&P500 stocks and non-S&P500 stocks, across the event dates. The betas of these portfolios with respect to the value-weighted market behave very differently. In results not tabulated here, we find that the average beta of the S&P500 portfolio decreases by 0.06 while that of the non-

S&P500 portfolio increases by 0.16. Depending on the other parameters, this effect alone would suggest an increase in the beta of a stock on the S&P500 of more than 0.20 relative to that on a portfolio of non-S&P500 stocks as shown in Section 5.3. This relative increase shows up primarily as an increase in beta on the S&P500 because the fundamental betas of the stocks in both the S&P500 addition and matched samples are also increasing. We speculate that the movements in the fundamental betas of the group portfolios are due to the technology boom at that time. As high risk technology stocks become more important in the overall market, the S&P500, which is relatively light in these stocks, exhibits a declining beta throughout this period. Regardless of the precise explanation, the fact that the effect shows up in the matched sample is clear evidence that it is a result of parameter instability at the group level.

For the bivariate regressions, the same basic results of no excess comovement hold. There are no statistically significant differences between the beta changes associated with the S&P addition sample and the matched sample. Moreover, while some of the individual beta changes have magnitudes of 0.1 or slightly higher in both samples, none of these individual differences is statistically significant. Again, the fact that similar patterns show up in the matched sample is an indication it is the properties of the group returns not the stocks that is changing across the event. In this case, the shifts in loadings across the two groups are consistent with changes in the relative fractions of idiosyncratic risk as illustrated in Section 5.2.

For stock splits, we have already established that even the original sample exhibits little or no evidence of excess comovement when comparing the univariate regression results across low-priced stocks (the new group) and high-priced stocks (the old group). Nevertheless, it is still worthwhile looking briefly at the results with Dimson betas for a momentum matched sample. Though we estimate Dimson betas for uniformity, we don't anticipate Dimson betas making a significant difference because non-synchronous trading is unlikely to be different for the high-priced and low-priced groups. On the other hand, almost all splitting stocks are likely to be momentum stocks so a properly matched sample should also exhibit similarly high changes in betas.

The basic results in Table 7 are affected little by the Dimson adjustment—comovement with both portfolios increases after the split by similar amounts. Not surprisingly, the same phenomenon shows up in the matched sample, although it is smaller than in the original sample. We attribute these differences to our inability to match some of the high returns on the splitting stocks in our matched sample. When taking differences across the samples, the values are economically very small and predominantly statistically insignificant.

Similar results obtain for the bivariate regressions and excess comovement is not evident in any sub-period except during 1991-2004. We attribute this result to the imperfect match. Nevertheless, the bivariate regression results are still puzzling. As an example, consider the results for both samples (i.e., the stock split sample and the matched sample) over the full period. In both cases, the coefficient on high priced stocks decreases, while that on low-priced stocks increases, and the changes are statistically significant in all cases. Clearly this result is not due to excess comovement since it shows up in the matched sample, and there is no change in group membership for these stocks. However, as noted in Section 5.2, small changes in the characteristics of the group portfolios can have large effects on these bivariate coefficients. In particular, increases in the idiosyncratic volatility of the returns on the high-priced group relative to that of the low-priced group are consistent with this phenomenon. A relative increase in idiosyncratic risk makes the group return a poorer proxy for the common (fundamental) factor, thus decreasing the weight that the regression puts on this return and increasing the weight on the other group return.

These results highlight the dangers of interpreting the coefficients from bivariate regressions, but they only strengthen our overall conclusion that there is no meaningful evidence of excess comovement.

7. Robustness Checks

We reconfirm the baseline results on comovement by repeating our analysis with weekly data and for index deletions.

7.1 Weekly Data

Though Barberis, Shleifer, and Wurgler (2005) and Green and Hwang (2009) present evidence of comovement using daily, weekly and monthly data, their results are strongest with daily data. Accordingly, the main results in the paper are based on daily data. Here we test the results with weekly data for S&P 500 additions and stock splits. Essentially, using weekly data has an effect similar to adding two leads and two lags to the beta estimates, as we do above. Not surprisingly, the results are much weaker with weekly data than with daily data. Once a matched sample is used to control for changes in fundamental factors, there is no evidence of residual excess comovement in univariate regressions for S&P 500 additions or for stock splits. In addition, there is no evidence of comovement for stock splits in the bivariate regressions.[14] However, there is weak evidence of comovement in bivariate regressions for the S&P 500 additions sample, which is not surprising given the prior discussion of instability of coefficients in bivariate regressions.

7.2 Index Deletions

Our baseline analysis has considered only S&P 500 index additions because they are more interesting, important, and the focus of prior research. Since stocks are both added to and deleted from the S&P500 index, usually at the same time, it is informative to also study index deletions for a reverse comovement effect. Unlike index additions, which are always voluntary and at the discretion of the Index committee of Standard and Poor's, index deletions may be voluntary or involuntary. Index deletions are involuntary when a firm ceases to exist (mergers and bankruptcies) or when a firm ceases to meet primary criteria established by Standard and Poor's (reincorporation in a foreign country). Voluntary index deletions may occur because a firm is no longer representative of the U.S. economy, the industry is less representative of the economy, or the firm has become too small in size.

[14] Results are not reported in the interests of brevity.

We repeat the analysis of S&P 500 index additions with a sample of primarily voluntary deletions.[15] Due to a smaller deletions sample and deleted firms potentially undergoing structural changes, we expect evidence of comovement in index deletions to be weaker. In addition, we anticipate that a significant fraction of the comovement may be explained by non-synchronous trading. We duplicate the analyses in Tables 1 and 6 for index deletions,[16] and find that the results are consistent with our results for index additions. Relative to Table 1, we find that the deleted firms move less with the S&P 500 index after deletion based on both univariate and bivariate regression coefficients, and that the results are primarily derived from the 1979-2000 period, as in Barberis, Shleifer, and Wurgler (2005). Relative to Table 6 with a matched sample and Dimson adjustments, we find that there is no residual evidence of excess comovement for the S&P 500 deletions sample. Thus, the analysis for index deletions corroborates evidence for index additions to suggest an absence of excess comovement.

8. Conclusion

Motivated by a simple model that captures the essence of the excess comovement hypothesis, we revisit the results of two well-known papers in the literature on comovement before and after S&P500 index additions (Barberis, Shleifer, and Wurgler, 2005) and stock splits (Green and Hwang, 2009). The model implies that looking at univariate regressions rather than bivariate regressions is more informative about the economic magnitude of the effect of interest, and, in particular, that the differences between the coefficients in univariate regressions on the returns of the group that the stock is leaving and the group that it is joining identify this effect. When we conduct this empirical exercise, the evidence points strongly to the conclusion that the existing results are due not to excess comovement but to changes in the comovement of stocks with fundamentals. These beta changes themselves are a feature common to winner stocks, an empirical phenomenon the documentation of which may be new to the literature. By making sure to measure these fundamental betas accurately, and controlling for this effect using a

[15] As for index additions, we extend the sample to 2012. The sample sizes for overlapping periods are similar to those in Barberis, Shleifer, and Wurgler (2005).
[16] Results are not tabulated here for brevity.

matched sample of winner stocks, we show that there is no longer any evidence of meaningful excess comovement from either an economic or statistical standpoint.

Appendix: Proofs

Assume the driving processes for returns prior to the group switch are

$$\begin{aligned} y_t &= \underline{b}_y f_t + \underline{c}_1 u_{1t} + e_{yt} \quad \underline{c}_1 > 0 \\ x_{1t} &= \underline{b}_1 f_t + u_{1t} + e_{1t} \\ x_{2t} &= \underline{b}_2 f_t + u_{2t} + e_{2t} \\ \operatorname{var}(e_{it}) &\equiv \underline{\sigma}_{ei}^2 \quad \operatorname{var}(u_{it}) \equiv \underline{\sigma}_{ui}^2 \quad \operatorname{var}(f) \equiv \underline{\sigma}_f^2 \end{aligned}$$

and similarly after the group switch

$$\begin{aligned} y_t &= \bar{b}_y f_t + \bar{c}_2 u_{2t} + e_{yt} \quad \bar{c}_2 > 0 \\ x_{1t} &= \bar{b}_1 f_t + u_{1t} + e_{1t} \\ x_{2t} &= \bar{b}_2 f_t + u_{2t} + e_{2t} \\ \operatorname{var}(e_{it}) &\equiv \bar{\sigma}_{ei}^2 \quad \operatorname{var}(u_{it}) \equiv \bar{\sigma}_{ui}^2 \quad \operatorname{var}(f) \equiv \bar{\sigma}_f^2 \end{aligned}$$

Univariate Regressions

In the univariate regressions

$$\begin{aligned} y_t &= \alpha + \beta_1 x_{1t} + \varepsilon_t \\ y_t &= \alpha + \beta_2 x_{2t} + \varepsilon_t \end{aligned},$$

the probability limit of the slope coefficient estimates are

$$\beta_1 = \frac{\operatorname{cov}(y_t, x_{1t})}{\operatorname{var}(x_{1t})} \quad \beta_2 = \frac{\operatorname{cov}(y_t, x_{2t})}{\operatorname{var}(x_{2t})}.$$

Computing the coefficients prior to and after the switch of stock *y* from group 1 to group 2:

$$\begin{aligned} \underline{\beta}_1 &= \frac{\operatorname{cov}(\underline{b}_y f_t + \underline{c}_1 u_{1t} + e_{yt}, \underline{b}_1 f_t + u_{1t} + e_{1t})}{\operatorname{var}(\underline{b}_1 f_t + u_{1t} + e_{1t})} = \frac{\underline{b}_y \underline{b}_1 \underline{\sigma}_f^2 + \underline{c}_1 \underline{\sigma}_{u1}^2}{\underline{\sigma}_{x1}^2} \\ \bar{\beta}_1 &= \frac{\operatorname{cov}(\bar{b}_y f_t + \bar{c}_2 u_{2t} + e_{yt}, \bar{b}_1 f_t + u_{1t} + e_{1t})}{\operatorname{var}(\bar{b}_1 f_t + u_{1t} + e_{1t})} = \frac{\bar{b}_y \bar{b}_1 \bar{\sigma}_f^2}{\bar{\sigma}_{x1}^2} \\ \underline{\sigma}_{x1}^2 &= \underline{b}_1^2 \underline{\sigma}_f^2 + \underline{\sigma}_{u1}^2 + \underline{\sigma}_{e1}^2 \quad \bar{\sigma}_{x1}^2 = \bar{b}_1^2 \bar{\sigma}_f^2 + \bar{\sigma}_{u1}^2 + \bar{\sigma}_{e1}^2 \end{aligned}.$$

Similarly,

$$\begin{aligned} \underline{\beta}_2 &= \frac{\underline{b}_y \underline{b}_2 \underline{\sigma}_f^2}{\underline{\sigma}_{x2}^2} \quad \bar{\beta}_2 = \frac{\bar{b}_y \bar{b}_2 \bar{\sigma}_f^2 + \bar{c}_2 \bar{\sigma}_{u2}^2}{\bar{\sigma}_{x2}^2} \\ \underline{\sigma}_{x2}^2 &= \underline{b}_2^2 \underline{\sigma}_f^2 + \underline{\sigma}_{u2}^2 + \underline{\sigma}_{e2}^2 \quad \bar{\sigma}_{x2}^2 = \bar{b}_2^2 \bar{\sigma}_f^2 + \bar{\sigma}_{u2}^2 + \bar{\sigma}_{e2}^2 \end{aligned}$$

Assuming the parameters other than stock *y*'s loadings on the fundamental factor and the non-fundamental group shocks are fixed across the 2 sub-periods, i.e.,

$$\underline{b}_1 = \overline{b}_1 \equiv b_1 \quad \underline{b}_2 = \overline{b}_2 \equiv b_2 \quad \underline{\sigma}_f^2 = \overline{\sigma}_f^2 \equiv \sigma_f^2 \quad \underline{\sigma}_{ui}^2 = \overline{\sigma}_{ui}^2 \equiv \sigma_{ui}^2 \quad \underline{\sigma}_{ei}^2 = \overline{\sigma}_{ei}^2 \equiv \sigma_{ei}^2$$

then

$$\overline{\beta}_1 - \underline{\beta}_1 = \frac{(\overline{b}_y - \underline{b}_y)\sigma_f^2 - \underline{c}_1\sigma_{u1}^2}{\sigma_{x1}^2}$$

$$\overline{\beta}_2 - \underline{\beta}_2 = \frac{(\overline{b}_y - \underline{b}_y)\sigma_f^2 + \overline{c}_2\sigma_{u2}^2}{\sigma_{x2}^2}$$

If, in addition $\underline{b}_y = \overline{b}_y \equiv b_y$, then

$$\overline{\beta}_1 - \underline{\beta}_1 = -\frac{\underline{c}_1\sigma_{u1}^2}{\sigma_{x1}^2} < 0$$

$$\overline{\beta}_2 - \underline{\beta}_2 = \frac{\overline{c}_2\sigma_{u2}^2}{\sigma_{x2}^2} > 0$$

Bivariate Regressions

Consider the bivariate regression:

$$y_t = \alpha + \beta_{1b} x_{1t} + \beta_{2b} x_{2t} + \varepsilon_t$$

The probability limits of the coefficients are

$$\beta = (X^T X)^{-1}(X^T Y) \quad \Rightarrow$$

$$\beta_{1b} = \frac{\operatorname{cov}(y_t, x_{1t})\operatorname{var}(x_{2t}) - \operatorname{cov}(y_t, x_{2t})\operatorname{cov}(x_{1t}, x_{2t})}{\operatorname{var}(x_{1t})\operatorname{var}(x_{2t}) - \operatorname{cov}(x_{1t}, x_{2t})^2}$$

$$\beta_{2b} = \frac{\operatorname{cov}(y_t, x_{2t})\operatorname{var}(x_{1t}) - \operatorname{cov}(y_t, x_{1t})\operatorname{cov}(x_{1t}, x_{2t})}{\operatorname{var}(x_{1t})\operatorname{var}(x_{2t}) - \operatorname{cov}(x_{1t}, x_{2t})^2}$$

where the coefficients reflect a natural symmetry. It is convenient to rewrite these expressions in terms of the univariate coefficients defined above:

$$\beta_{1b} = \frac{\operatorname{cov}(y_t, x_{1t})\operatorname{var}(x_{2t}) - \operatorname{cov}(y_t, x_{2t})\operatorname{corr}(x_{1t}, x_{2t})\sqrt{\operatorname{var}(x_{1t})\operatorname{var}(x_{2t})}}{(1 - \operatorname{corr}(x_{1t}, x_{2t})^2)\operatorname{var}(x_{1t})\operatorname{var}(x_{2t})}$$

$$= \frac{1}{1 - \operatorname{corr}(x_{1t}, x_{2t})^2}\left(\frac{\operatorname{cov}(y_t, x_{1t})}{\operatorname{var}(x_{1t})} - \operatorname{corr}(x_{1t}, x_{2t})\sqrt{\frac{\operatorname{var}(x_{2t})}{\operatorname{var}(x_{1t})}}\frac{\operatorname{cov}(y_t, x_{2t})}{\operatorname{var}(x_{2t})}\right)$$

$$= \frac{1}{1 - \operatorname{corr}(x_{1t}, x_{2t})^2}\left(\beta_1 - \operatorname{corr}(x_{1t}, x_{2t})\sqrt{\frac{\operatorname{var}(x_{2t})}{\operatorname{var}(x_{1t})}}\beta_2\right)$$

$$\beta_{2b} = \frac{1}{1 - \operatorname{corr}(x_{1t}, x_{2t})^2}\left(\beta_2 - \operatorname{corr}(x_{1t}, x_{2t})\sqrt{\frac{\operatorname{var}(x_{1t})}{\operatorname{var}(x_{2t})}}\beta_1\right)$$

As above, computing these values prior to and after the switch of stock *y* from group 1 to group 2:

$$\underline{\beta}_{1b} = \frac{1}{1 - \underline{\rho}_{x1,x2}^2}\left[\underline{\beta}_1 - \underline{\rho}_{x1,x2}\frac{\underline{\sigma}_{x2}}{\underline{\sigma}_{x1}}\underline{\beta}_2\right] \quad \underline{\beta}_{2b} = \frac{1}{1 - \underline{\rho}_{x1,x2}^2}\left[\underline{\beta}_2 - \underline{\rho}_{x1,x2}\frac{\underline{\sigma}_{x1}}{\underline{\sigma}_{x2}}\underline{\beta}_1\right]$$

$$\bar{\beta}_{1b} = \frac{1}{1 - \bar{\rho}_{x1,x2}^2}\left[\bar{\beta}_1 - \bar{\rho}_{x1,x2}\frac{\bar{\sigma}_{x2}}{\bar{\sigma}_{x1}}\bar{\beta}_2\right] \quad \bar{\beta}_{2b} = \frac{1}{1 - \bar{\rho}_{x1,x2}^2}\left[\bar{\beta}_2 - \bar{\rho}_{x1,x2}\frac{\bar{\sigma}_{x1}}{\bar{\sigma}_{x2}}\bar{\beta}_1\right]$$

$$\underline{\rho}_{x1,x2} = \frac{\operatorname{cov}(\underline{x}_1, \underline{x}_2)}{\underline{\sigma}_{x1}\underline{\sigma}_{x2}} \quad \operatorname{cov}(\underline{x}_1, \underline{x}_2) = \underline{b}_1\underline{b}_2\underline{\sigma}_f^2 + \operatorname{cov}(\underline{e}_1, \underline{e}_2)$$

$$\bar{\rho}_{x1,x2} = \frac{\operatorname{cov}(\bar{x}_1, \bar{x}_2)}{\bar{\sigma}_{x1}\bar{\sigma}_{x2}} \quad \operatorname{cov}(\bar{x}_1, \bar{x}_2) = \bar{b}_1\bar{b}_2\bar{\sigma}_f^2 + \operatorname{cov}(\bar{e}_1, \bar{e}_2)$$

Again assuming the parameters other than the weights on the non-fundamental group shocks are fixed across the 2 sub-periods,

$$\underline{\beta}_{1b} - \bar{\beta}_{1b} = \frac{1}{1 - \rho_{x1,x2}^2}\left[\left(\underline{\beta}_1 - \bar{\beta}_1\right) - \rho_{x1,x2}\frac{\sigma_{x2}}{\sigma_{x1}}\left(\underline{\beta}_2 - \bar{\beta}_2\right)\right] > 0$$

$$\underline{\beta}_{2b} - \bar{\beta}_{2b} = \frac{1}{1 - \rho_{x1,x2}^2}\left[\left(\underline{\beta}_2 - \bar{\beta}_2\right) - \rho_{x1,x2}\frac{\sigma_{x1}}{\sigma_{x2}}\left(\underline{\beta}_1 - \bar{\beta}_1\right)\right] < 0$$

If we further assume

$$\sigma_{e1}^2 = \sigma_{e2}^2 = 0 \quad \underline{c}_1 = \bar{c}_2 = 1 \quad b_y = b_1 = b_2 = 1$$

then

$$\underline{\beta}_1 = \frac{\sigma_f^2 + \sigma_{u1}^2}{\sigma_f^2 + \sigma_{u1}^2} = 1 \quad \overline{\beta}_1 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_{u1}^2}$$

$$\underline{\beta}_2 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_{u2}^2} \quad \overline{\beta}_2 = \frac{\sigma_f^2 + \sigma_{u2}^2}{\sigma_f^2 + \sigma_{u2}^2} = 1$$

$$\rho_{x1,x2} = \frac{\sigma_f^2}{\sqrt{(\sigma_f^2 + \sigma_{u1}^2)(\sigma_f^2 + \sigma_{u2}^2)}}$$

$$\underline{\beta}_{1b} = \frac{1}{1-\rho_{x1,x2}^2}\left[1 - \rho_{x1,x2}\sqrt{\frac{\sigma_f^2 + \sigma_{u2}^2}{\sigma_f^2 + \sigma_{u1}^2}}\,\frac{\sigma_f^2}{\sigma_f^2 + \sigma_{u2}^2}\right] = \frac{1}{1-\rho_{x1,x2}^2}\left[1 - \rho_{x1,x2}^2\right] = 1$$

$$\underline{\beta}_{2b} = \frac{1}{1-\rho_{x1,x2}^2}\left[\frac{\sigma_f^2}{\sigma_f^2 + \sigma_{u2}^2} - \frac{\sigma_f^2}{\sqrt{(\sigma_f^2 + \sigma_{u1}^2)(\sigma_f^2 + \sigma_{u2}^2)}}\sqrt{\frac{\sigma_f^2 + \sigma_{u1}^2}{\sigma_f^2 + \sigma_{u2}^2}}\right] = 0$$

Similarly,

$$\overline{\beta}_{1b} = 0 \quad \overline{\beta}_{2b} = 1$$

References

Barberis, N., & Shleifer, A., 2003. Style investing. Journal of Financial Economics, 68.

Barberis, N., Shleifer, A., & Wurgler, J., 2005. Comovement. Journal of Financial Economics, 75(2), 283–317.

Basak, S., & Pavlova, A., 2013. Asset prices and institutional investors. American Economic Review, 103(5), 1728-1758.

Bodurtha, J. N., Kim, D. S., & Lee, C. M., 1995. Closed-end country funds and US market sentiment. Review of Financial Studies, 8(3), 879-918.

Boyer, B.H., 2011. Style-related comovement: Fundamentals or labels? Journal of Finance, 66(1), 307–32.

Chen, H., Noronha, G. & Singal, V., 2004, The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation, Journal of Finance, 59(4), 1901-1930.

Claessens, S., & Yafeh, Y., 2011. Additions to market indices and the comovement of stock returns around the world. IMF Working Paper, #11/47.

DeMarzo, P.M., Kaniel, R., & Kremer, I., 2004. Diversification as a public good: Community effects in portfolio choice. Journal of Finance, 59(4), 1677–1715.

Denis, D. J., & G. B. Kadlec, 1994, Corporate events, trading activity, and the estimation of systematic risk: Evidence form equity offerings and share repurchases, Journal of Finance, 49(5), 1787–1811.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics, 7(2), 197-226.

Elliott, W.B., Van Ness, B.F., Walker, M.D., & Wan, R.S., 2006. What drives the S&P 500 inclusion effect? An analytical survey. Financial Management, 35(4), 31-48.

Green, T.C., & Hwang, B., 2009. Price-based return comovement. Journal of Financial Economics 93, 37–50.

Greenwood, R., & Sosner, N., 2007. Trading patterns and excess comovement of stock returns. Financial Analysts Journal, 63, 69-81.

Greenwood, R., 2008, Excess comovement and stock returns; Evidence from cross-sectional variation in Nikkei 225 weights. Review of Financial Studies, 21, 1153-186.

Hardouvelis, G., Porta, R.L., & Wizman, T.A., 1994. What moves the discount on country equity funds?. In The internationalization of equity markets (pp. 345-403). University of Chicago Press.

Hou, K., & Moskowitz, T.J., 2005. Market frictions, price delay, and the cross-section of expected returns. Review of Financial Studies, 18(3), 981-1020.

Jegadeesh, N., & Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. Journal of Finance, 56, 699–720.

Kasch, M., & Sarkar, A., 2012, Is there an S&P500 index effect? Federal Reserve Bank Working paper, SSRN 2171235.

Lee, M. C., Shleifer, A., & Thaler, R.H., 1991, Investment sentiment and the closed-end fund puzzle, Journal of Finance 46, 76–110.

Mase, B., 2008. Comovement in the FTSE 100 index. Applied Financial Economics Letters, 4, 9-12.

Pindyck, R., & Rotemberg, J., 1993. The comovement of stock prices. Quarterly Journal of Economics, 108, 1073-104.

Pirinsky, C., & Wang, Q., 2006. Does corporate headquarters location matter for stock returns? Journal of Finance, 61(4), 1991-2015.

Rigobon, R., 2002. The curse of non-investment grade countries. Journal of Development Economics, 69(2), 423–49.

Spanos, A., & McGuirk, A., 2002. The problem of near-multicollinearity revisited: Erratic vs. systematic volatility. Journal of Econometrics, 108(2), 365-393.

Tang, K. & Xiong, W., 2012, Index investing and the financialization of commodities, Financial Analysts Journal, 68(6): 54-74.

Vijh, A., 1994. S&P trading strategies and stock betas. Review of Financial Studies, 7, 215-51.

Von Drathen, C., 2013. Is there really excess comovement? Causal evidence from FTSE 100 index turnover, Working paper, LSE.

Table 1: S&P Additions

We estimate the univariate and bivariate regressions

$$y_t = \alpha + \beta_1 x_{1t} + \varepsilon_t$$
$$y_t = \alpha + \beta_2 x_{2t} + \varepsilon_t$$
$$y_t = \alpha + \beta_{1b} x_{1t} + \beta_{2b} x_{2t} + \varepsilon_t$$

for a sample of stocks that are added to the S&P 500 index from 1962 through 2012. The pre-event estimation period covers a one year window ending at the end of the month preceding announcement, while the post-event period covers the one year window starting the month after the effective date of index change. x_{1t} and x_{2t} are returns to non-S&P 500 index and S&P 500 index at time t. Panel A reports the univariate regression results, and Panel B reports the bivariate regression results. In each cell, the first number is the mean and the second number is the corresponding t-statistic, where standard errors are clustered by month.

Panel A. Univariate Regressions

		Non-S&P500 Group			S&P500			Diff. of Diff.
Sample Period	nobs	$\underline{\beta}_1$	$\overline{\beta}_1$	$\Delta\beta_1$	$\underline{\beta}_2$	$\overline{\beta}_2$	$\Delta\beta_2$	$\Delta\beta_2 - \Delta\beta_1$
1976-1987	197	1.271 29.422	1.313 27.202	0.042 1.055	0.962 24.643	1.024 26.830	0.062 2.305	0.020 0.763
1988-2000	269	1.263 29.839	1.278 27.503	0.015 0.313	0.984 24.669	1.198 23.830	0.214 6.243	0.199 4.938
2001-2012	214	1.050 30.548	1.125 28.736	0.075 2.496	1.086 27.526	1.157 35.741	0.071 2.439	-0.004 -0.137
1976-2012	680	1.198 49.851	1.240 46.610	0.042 1.734	1.010 43.566	1.134 44.294	0.125 6.556	0.083 4.080

Panel B. Bivariate Regressions

		Non-S&P Group			S&P500			Diff. of Diff.
Sample Period	nobs	$\underline{\beta}_{1b}$	$\overline{\beta}_{1b}$	$\Delta\beta_{1b}$	$\underline{\beta}_{2b}$	$\overline{\beta}_{2b}$	$\Delta\beta_{2b}$	$\Delta\beta_{2b} - \Delta\beta_{1b}$
1976-1987	197	0.907 15.366	0.632 11.692	-0.274 -4.627	0.340 5.884	0.602 10.626	0.262 5.733	0.537 5.377
1988-2000	269	1.011 23.314	0.647 11.554	-0.364 -5.877	0.281 10.557	0.667 18.177	0.386 8.789	0.750 7.373
2001-2012	214	0.951 13.407	0.691 9.456	-0.260 -4.323	0.127 2.229	0.473 9.323	0.347 6.379	0.607 5.414
1976-2012	680	0.962 29.394	0.657 18.504	-0.305 -8.708	0.249 9.143	0.587 21.347	0.338 12.249	0.643 10.675

Table 2: Stock Splits

We estimate the univariate and bivariate regressions

$$y_t = \alpha + \beta_1 x_{1t} + \varepsilon_t ; y_t = \alpha + \beta_2 x_{2t} + \varepsilon_t$$

$$y_t = \alpha + \beta_{1b} x_{1t} + \beta_{2b} x_{2t} + \varepsilon_t$$

for a sample of 2-for-1 stock splits from 1962 through 2012. Our sample include all ordinary common stock two-for-one splits with a pre-split price of $10 or greater during our sample period. x_{1t} and x_{2t} are return to a portfolio of high priced stocks whose price belongs to [3p/4, 5p/4] and low price stocks with prices within [1p/4, 3p/4] at time t, where p is the pre-split price before effective date of split. The pre-event (post-event) window is defined as the one year ending (beginning) one month before (after) the split date. Panel A reports the univariate regression results and Panel B reports the bivariate regression results. In each cell, the first number is the mean and the second number is the corresponding t-statistic, where standard errors are clustered by month.

Panel A. Univariate Regressions

		High-Priced Group			Low-Priced Group			Diff. of Diff.
Sample Period	nobs	$\underline{\beta}_1$	$\overline{\beta}_1$	$\Delta\beta_1$	$\underline{\beta}_2$	$\overline{\beta}_2$	$\Delta\beta_2$	$\Delta\beta_2 - \Delta\beta_1$
1971-1990	2,350	0.736 48.771	0.929 54.709	0.193 17.138	0.847 49.059	1.043 60.555	0.196 18.554	0.002 0.443
1991-2004	2,478	0.798 45.714	1.014 43.778	0.216 11.102	0.937 43.783	1.186 39.479	0.248 12.020	0.033 3.448
1976-1987	1,867	0.729 40.663	0.919 45.952	0.190 14.620	0.847 40.182	1.036 50.304	0.189 15.136	-0.001 -0.189
1988-2000	2,383	0.796 44.584	1.001 42.256	0.205 10.371	0.968 46.228	1.206 39.782	0.237 11.141	0.032 3.256
2001-2012	794	0.932 30.602	1.141 36.926	0.209 8.553	0.887 28.309	1.084 36.203	0.197 8.518	-0.012 -0.976
1976-2012	5,044	0.792 63.604	0.993 66.954	0.200 17.930	0.910 64.844	1.124 63.436	0.213 18.142	0.013 2.281

Panel B. Bivariate Regressions

		High-Priced Group			Low-Priced Group			Diff. of Diff.
Sample Period	nobs	$\underline{\beta}_{1b}$	$\overline{\beta}_{1b}$	$\Delta\beta_{1b}$	$\underline{\beta}_{2b}$	$\overline{\beta}_{2b}$	$\Delta\beta_{2b}$	$\Delta\beta_{2b}-\Delta\beta_{1b}$
1971-1990	2,350	-0.013	-0.043	-0.030	0.865	1.085	0.220	0.250
		-0.724	-1.721	-1.011	38.366	41.082	7.235	4.231
1991-2004	2,478	0.041	0.003	-0.038	0.883	1.171	0.289	0.326
		1.379	0.092	-1.271	28.630	34.457	6.884	4.729
1976-1987	1,867	-0.016	-0.068	-0.052	0.863	1.101	0.238	0.290
		-0.938	-2.336	-1.539	34.684	35.398	6.884	4.311
1988-2000	2,383	0.001	-0.035	-0.036	0.951	1.224	0.273	0.309
		0.027	-1.096	-1.160	29.898	37.527	6.124	4.229
2001-2012	794	0.356	0.420	0.064	0.568	0.707	0.140	0.075
		6.585	8.493	1.594	14.599	17.584	3.841	1.036
1976-2012	5,044	0.050	0.024	-0.026	0.858	1.097	0.239	0.265
		2.598	1.085	-1.271	43.961	48.520	9.363	5.956

Table 3: Numerical Examples

We calculate the univariate and bivariate regression coefficients implied by the model in Section 2 for four numerical examples. In all cases, we assume no excess comovement, no correlation between idiosyncratic shocks at the group level, the same amount of idiosyncratic risk at the stock level, and the same amount of fundamental risk:

$$\underline{\sigma}_{u1}^2 = \bar{\sigma}_{u1}^2 = \underline{\sigma}_{u2}^2 = \bar{\sigma}_{u2}^2 = 0 \quad \mathrm{cov}(\underline{e}_1,\underline{e}_2) = \mathrm{cov}(\bar{e}_1,\bar{e}_2) = 0 \quad \underline{\sigma}_{ey} = \bar{\sigma}_{ey} = 1.73\% \quad \underline{\sigma}_f = \bar{\sigma}_f = 1\%$$

The base case (top row) assumes perfect symmetry. The subsequent examples allow for parameter instability across the event, specifically (1) a change in the stock beta, (2) a change in idiosyncratic risk at the group level, and (3) a change in group beta. In each case, the deviations from the base case for both the input parameters and the regression coefficients are highlighted in bold.

Panel A. Inputs

Case types	Fundamental Loadings						Group Idiosyncratic Volatility			
	$\underline{b}_1$	$\bar{b}_1$	$\underline{b}_2$	$\bar{b}_2$	$\underline{b}_y$	$\bar{b}_y$	$\underline{\sigma}_{e1}$	$\bar{\sigma}_{e1}$	$\underline{\sigma}_{e2}$	$\bar{\sigma}_{e2}$
Base Case	1.0	1.0	1.0	1.0	1.0	1.0	0.20%	0.20%	0.20%	0.20%
(1) Change in stock beta	1.0	1.0	1.0	1.0	1.0	**1.2**	0.20%	0.20%	0.20%	0.20%
(2) Change in group i-risk	1.0	1.0	1.0	1.0	1.0	1.0	0.20%	**0.24%**	0.20%	0.20%
(3) Change in group beta	1.0	1.0	1.0	**0.8**	1.0	1.0	0.20%	0.20%	0.20%	0.20%

Panel B. Regression Coefficients

	Univariate						Bivariate					
Case types	Coefficient				Change		Coefficient				Change	
	$\underline{\beta}_1$	$\bar{\beta}_1$	$\underline{\beta}_2$	$\bar{\beta}_2$	$\bar{\beta}_1 - \underline{\beta}_1$	$\bar{\beta}_2 - \underline{\beta}_2$	$\underline{\beta}_{1b}$	$\bar{\beta}_{1b}$	$\underline{\beta}_{2b}$	$\bar{\beta}_{2b}$	$\bar{\beta}_{1b}$ - $\underline{\beta}_{1b}$	$\bar{\beta}_{2b}$ - $\underline{\beta}_{2b}$
Base Case	0.962	0.962	0.962	0.962	0.000	0.000	0.490	0.490	0.490	0.490	0.000	0.000
(1) Change in stock beta	0.962	**1.154**	0.962	**1.154**	**0.192**	**0.192**	0.490	**0.588**	0.490	**0.588**	**0.098**	**0.098**
(2) Change in group i-risk	0.962	**0.946**	0.962	0.962	**-0.016**	0.000	0.490	**0.400**	0.490	**0.577**	**-0.090**	**0.086**
(3) Change in group beta	0.962	0.962	0.962	**1.176**	0.000	**0.215**	0.490	**0.595**	0.490	**0.476**	**0.105**	**-0.014**

Table 4: Past Return Performance of Sample Stocks

For each stock in the sample we record which momentum decile portfolio it would be in based on its returns over the prior 12 months. The table reports the percentage of stocks in the S&P500 additions and stock splits samples that fall in each decile and the mean and median return on these stocks over the prior year.

	Frequency (%)	
Decile	S&P500 Additions	Stock Splits
Losers	2.94	0.28
2	3.68	0.89
3	5.00	1.61
4	4.12	2.92
5	12.06	4.78
6	9.12	7.66
7	11.91	10.09
8	14.12	14.75
9	16.62	21.36
Winners	20.44	35.66
Mean Return	41.6%	109.1%
Median Return	25.0%	63.9%

Table 5: Beta Changes and Momentum

At the end of each June from 1976 through 2011, stocks with a price of at least $10 that do not fall into the bottom size decile of NYSE stocks are assigned into 10 momentum deciles based on their cumulative returns over the preceding 252 days. We estimate betas for each stock based on a rolling window of 252 days from two years before formation of momentum portfolios through two years after formation, and compare beta changes for both the top and bottom two momentum portfolios. Thus, betas for years -2 and -1 are estimated over rolling windows ending 504 and 252 trading days before portfolio formation, respectively. Post-momentum portfolio formation years allow for a 21-trading day skip, and are estimated over 252 days ending 273 and 525 trading days after portfolio formation. In each cell, the first number is the time series average of the mean, and the second number is the corresponding t-statistic.

Momentum Decile	Year -2	Year -1	Year 0	Year 1	Year 2	Year 0 – Year -2	Year 0 – Year -1	Year 1 – Year 0	Year 2 – Year 0
10 (Winners)	0.964	0.976	1.143	1.271	1.166	0.179	0.167	0.128	0.023
	21.865	20.588	21.313	21.163	25.997	4.139	4.891	3.217	0.553
9	0.918	0.916	0.981	1.038	0.994	0.063	0.065	0.057	0.013
	21.974	21.109	21.917	23.529	27.510	2.161	2.693	2.194	0.448
Middle 6 Deciles	0.818	0.829	0.832	0.834	0.841	0.014	0.003	0.002	0.009
	29.606	28.510	27.432	26.003	25.132	0.642	0.221	0.150	0.438
2	0.903	0.925	0.917	0.876	0.888	0.014	-0.007	-0.042	-0.029
	28.188	28.429	23.377	20.214	20.436	0.579	-0.373	-2.546	-1.251
1 (Losers)	1.047	1.094	1.092	1.031	1.015	0.045	-0.003	-0.061	-0.077
	31.351	32.901	21.759	20.707	21.423	1.229	-0.091	-1.977	-2.219

Table 6: S&P Additions with Matched Sample and Dimson Adjustments

We estimate the univariate and bivariate Dimson (1979) regressions for a sample of stocks that are added to the S&P 500 index from 1976 through 2012 and for a portfolio of matched firms. The pre-event estimation period covers a one year window ending at the end of the month preceding announcement, while the post-event period covers the one year window starting the month after the effective date of index change. x_{1t} is return to non-S&P 500 index at time t, while x_{2t} is return to the S&P 500 index at time t. The match firm for each addition is identified as the one with closest momentum from the same size decile as the addition firms. The Dimson beta is defines as a simple sum of the lag, concurrent, and lead coefficients from the following regressions with two leads and lags. In each cell, the first number is the mean and the second number is the corresponding t-statistic, where standard errors are clustered by month.

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_1^s x_{1,t+s} + \varepsilon_t$$

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_2^s x_{2,t+s} + \varepsilon_t$$

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_1^s x_{1,t+s} + \sum_{s=-2}^{2} \beta_2^s x_{2,t+s} + \varepsilon_t$$

Panel A. Univariate Regressions

	Sample Period	nobs	Non-S&P500 Group			S&P500			Diff. of Diff.
			$\underline{\beta}_1$	$\overline{\beta}_1$	$\Delta\beta_1$	$\underline{\beta}_2$	$\overline{\beta}_2$	$\Delta\beta_2$	$\Delta\beta_2 - \Delta\beta_1$
Sample	1976-1987	187	1.190	1.272	0.081	1.156	1.190	0.033	-0.048
			28.734	28.294	2.017	26.699	29.071	0.820	-2.096
	1988-2000	245	1.193	1.142	-0.051	1.161	1.239	0.078	0.129
			27.716	23.419	-0.919	29.551	25.360	1.527	2.549
	2001-2012	203	1.007	1.020	0.013	1.187	1.168	-0.020	-0.033
			26.383	25.779	0.348	29.117	27.880	-0.483	-1.187
	1976-2012	635	1.133	1.141	0.008	1.168	1.202	0.034	0.025
			46.873	42.120	0.312	49.382	45.861	1.260	1.093
Match	1976-1987	187	1.060	1.066	0.006	0.985	0.989	0.004	-0.002
			27.537	26.633	0.143	27.133	23.089	0.103	-0.085
	1988-2000	245	1.103	1.071	-0.031	1.087	1.167	0.080	0.111
			23.047	19.611	-0.493	25.860	17.404	1.156	2.232
	2001-2012	203	0.976	0.994	0.017	1.138	1.132	-0.006	-0.023
			21.149	24.104	0.437	22.733	23.749	-0.126	-0.877
	1976-2012	635	1.050	1.045	-0.005	1.073	1.103	0.030	0.035
			40.310	38.185	-0.160	42.135	33.184	0.916	1.552
Sample -Match	1976-1987	187	0.130	0.206	0.076	0.171	0.201	0.030	-0.046
			2.721	3.783	1.321	3.965	3.752	0.550	-1.998
	1988-2000	245	0.090	0.070	-0.020	0.074	0.072	-0.002	0.018
			1.821	1.363	-0.392	1.481	1.445	-0.029	0.601
	2001-2012	203	0.031	0.026	-0.004	0.050	0.036	-0.014	-0.009
			0.537	0.506	-0.094	0.806	0.660	-0.247	-0.458
	1976-2012	635	0.083	0.096	0.013	0.095	0.099	0.004	-0.009
			2.764	3.097	0.447	3.104	3.179	0.121	-0.631

Panel B. Bivariate Regressions

	Sample Period	nobs	Non-S&P500 Group			S&P500			Diff. of Diff.
			$\underline{\beta}_{1b}$	$\overline{\beta}_{1b}$	$\Delta\beta_{1b}$	$\underline{\beta}_{2b}$	$\overline{\beta}_{2b}$	$\Delta\beta_{2b}$	$\Delta\beta_{2b}-\Delta\beta_{1b}$
Sample	1976-1987	187	0.756 10.061	0.809 10.040	0.053 0.489	0.462 7.552	0.452 5.819	-0.010 -0.100	-0.063 -0.307
	1988-2000	245	0.803 14.017	0.693 9.300	-0.110 -1.271	0.438 8.225	0.562 8.152	0.123 1.579	0.233 1.473
	2001-2012	203	0.901 10.262	0.791 7.895	-0.109 -1.220	0.148 1.765	0.270 2.817	0.123 1.219	0.232 1.241
	1976-2012	635	0.820 19.467	0.759 15.474	-0.062 -1.145	0.352 8.889	0.436 9.269	0.084 1.583	0.146 1.399
Match	1976-1987	187	0.844 12.157	0.803 12.962	-0.041 -0.491	0.221 3.462	0.265 4.522	0.044 0.558	0.085 0.542
	1988-2000	245	0.800 12.115	0.785 10.199	-0.015 -0.177	0.368 6.708	0.406 5.661	0.038 0.452	0.053 0.332
	2001-2012	203	0.893 8.893	0.808 7.155	-0.086 -0.819	0.107 1.087	0.209 1.918	0.102 0.905	0.188 0.881
	1976-2012	635	0.843 18.253	0.798 15.937	-0.045 -0.856	0.241 5.467	0.301 6.242	0.060 1.122	0.105 1.027
Sample -Match	1976-1987	187	-0.087 -0.816	0.006 0.059	0.093 0.648	0.241 2.806	0.187 1.886	-0.054 -0.415	-0.147 -0.550
	1988-2000	245	0.003 0.040	-0.092 -1.035	-0.095 -1.019	0.070 0.942	0.156 1.955	0.086 0.930	0.181 1.014
	2001-2012	203	0.008 0.067	-0.016 -0.111	-0.024 -0.188	0.041 0.392	0.061 0.432	0.020 0.147	0.044 0.170
	1976-2012	635	-0.022 -0.399	-0.039 -0.598	-0.017 -0.244	0.111 2.179	0.135 2.178	0.024 0.347	0.041 0.304

Table 7: Stock Splits with Matched Sample and Dimson Adjustments

We estimate the univariate and bivariate Dimson (1979) regressions for a sample of sample of 2-for-1 stock splits from 1976 through 2012. Our sample include all ordinary common stock two-for-one splits with a pre-split price of $10 or greater during our sample period. x_{1t} and x_{2t} are return to a portfolio of high priced stocks whose price belongs to [3p/4, 5p/4] and low price stocks with prices within [1p/4, 3p/4] at time t, where p is the pre-split price before effective date of split. The pre-event (post-event) window is defined as the one year ending (beginning) one month before (after) the split date. The Dimson beta is defines as a simple sum of the lag, concurrent, and lead coefficients from the following regressions with two leads and lags. In each cell, the first number is the mean and the second number is the corresponding t-statistic, where standard errors are clustered by month. In each cell, the first number is the mean and the second number is the corresponding t-statistic, where standard errors are clustered by month.

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_1^s x_{1,t+s} + \varepsilon_t$$

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_2^s x_{2,t+s} + \varepsilon_t$$

$$y_t = \alpha + \sum_{s=-2}^{2} \beta_1^s x_{1,t+s} + \sum_{s=-2}^{2} \beta_2^s x_{2,t+s} + \varepsilon_t$$

Panel A. Univariate Regressions

			High-Priced Group			Low-Priced Group			Diff. of Diff.
	Sample Period	nobs	$\underline{\beta}_1$	$\overline{\beta}_1$	$\Delta\beta_1$	$\underline{\beta}_2$	$\overline{\beta}_2$	$\Delta\beta_2$	$\Delta\beta_2 - \Delta\beta_1$
Sample	1976-1987	1,606	0.924 55.306	1.120 42.904	0.197 8.486	0.981 55.824	1.154 51.170	0.172 8.603	-0.024 -2.271
	1988-2000	2,097	0.963 42.614	1.132 41.875	0.168 5.572	1.056 45.824	1.267 40.457	0.211 7.029	0.043 2.834
	2001-2012	727	1.042 29.491	1.222 37.391	0.180 5.181	0.961 26.458	1.135 34.593	0.174 5.621	-0.006 -0.411
	1976-2012	4,430	0.962 69.502	1.142 67.815	0.181 10.295	1.014 72.336	1.204 66.694	0.191 11.367	0.010 1.172
Match	1976-1987	1,606	0.858 48.739	0.957 42.268	0.099 5.078	0.901 50.803	0.981 48.675	0.080 4.855	-0.019 -2.349
	1988-2000	2,097	0.847 46.451	0.943 40.139	0.095 3.492	0.926 49.136	1.046 33.341	0.120 4.236	0.025 2.112
	2001-2012	727	0.991 30.774	1.116 42.113	0.125 4.538	0.917 29.015	1.029 37.387	0.112 4.852	-0.013 -1.031
	1976-2012	4,430	0.875 70.200	0.976 66.242	0.102 6.620	0.915 75.350	1.020 59.093	0.104 6.856	0.003 0.375
Sample -Match	1976-1987	1,606	0.066 4.459	0.163 8.755	0.098 4.774	0.080 5.694	0.173 9.602	0.093 4.620	-0.005 -0.688
	1988-2000	2,097	0.116 6.757	0.189 9.134	0.073 3.493	0.131 7.490	0.222 10.703	0.091 4.638	0.018 1.619
	2001-2012	727	0.050 1.823	0.105 3.372	0.055 1.790	0.045 1.727	0.106 3.437	0.062 2.099	0.007 0.582
	1976-2012	4,430	0.087 8.081	0.166 12.819	0.079 5.911	0.098 9.204	0.185 14.350	0.087 6.826	0.008 1.249

Panel B. Bivariate Regressions

	Sample Period	nobs	High-Priced Group $\underline{\beta}_{1b}$	$\overline{\beta}_{1b}$	$\Delta\beta_{1b}$	Low-Priced Group $\underline{\beta}_{2b}$	$\overline{\beta}_{2b}$	$\Delta\beta_{2b}$	Diff. of Diff. $\Delta\beta_{2b}-\Delta\beta_{1b}$
Sample	1976-1987	1,606	0.262 7.176	0.152 3.754	-0.109 -1.915	0.722 20.178	0.993 25.398	0.272 4.645	0.381 3.349
	1988-2000	2,097	0.177 4.073	0.008 0.161	-0.170 -3.172	0.866 18.434	1.238 27.571	0.372 5.726	0.542 4.741
	2001-2012	727	0.471 7.325	0.464 8.411	-0.007 -0.101	0.554 11.453	0.731 13.473	0.176 3.055	0.183 1.513
	1976-2012	4,430	0.256 9.363	0.135 4.514	-0.121 -3.492	0.762 27.557	1.066 36.602	0.304 7.847	0.425 5.973
Match	1976-1987	1,606	0.317 9.210	0.172 4.637	-0.144 -2.832	0.588 17.657	0.810 20.767	0.221 4.132	0.366 3.549
	1988-2000	2,097	0.226 6.847	0.109 3.114	-0.117 -2.725	0.698 21.726	0.941 24.571	0.243 4.906	0.360 4.129
	2001-2012	727	0.455 8.797	0.482 9.691	0.027 0.432	0.518 12.138	0.616 14.502	0.098 1.893	0.072 0.651
	1976-2012	4,430	0.296 13.273	0.193 7.960	-0.103 -3.482	0.628 29.628	0.840 33.409	0.211 6.631	0.315 5.314
Sample-Match	1976-1987	1,606	-0.055 -1.279	-0.020 -0.508	0.035 0.603	0.133 3.086	0.184 4.745	0.050 0.841	0.015 0.134
	1988-2000	2,097	-0.049 -1.240	-0.101 -2.548	-0.053 -0.944	0.168 4.067	0.298 7.071	0.130 2.347	0.182 1.672
	2001-2012	727	0.015 0.231	-0.018 -0.268	-0.033 -0.470	0.037 0.604	0.115 1.745	0.078 1.116	0.112 0.813
	1976-2012	4,430	-0.040 -1.517	-0.058 -2.228	-0.018 -0.496	0.134 4.955	0.226 8.443	0.092 2.575	0.110 1.568

<u>Figure 1: Beta Changes and Momentum</u>

We estimate market betas of winner and loser stocks defined as the top and bottom deciles of stocks sorted on past 12-month returns, skipping the most recent month, as in Jegadeesh and Titman (2001), for the sample period 1976-2011. These betas are estimated over rolling windows of 252 days (1 year).