Pseudo Market Timing and Predictive Regressions*

Malcolm Baker
Harvard Business School and NBER
mbaker@hbs.edu

Ryan Taliaferro
Harvard Business School
rtaliaferro@hbs.edu

Jeffrey Wurgler
NYU Stern School of Business and NBER
jwurgler@stern.nyu.edu

September 28, 2004

Abstract

A number of studies claim that aggregate managerial decision variables, such as aggregate equity issuance, have power to predict stock or bond market returns. Recent research argues that these results may be driven by an aggregate time-series version of Schultz's (2003) pseudo market timing bias. We use standard simulation techniques to estimate the size of the aggregate pseudo market timing bias for a variety of predictive regressions based on managerial decision variables. We find that the bias can explain only about one percent of the predictive power of the equity share in new issues, and that it is also much too small to overturn prior inferences about the predictive power of corporate investment plans, insider trading, dividend initiations, or the maturity of corporate debt issues.

* We thank Yakov Amihud, John Campbell, Owen Lamont, Alex Ljungqvist, Tim Loughran, Jay Ritter, Andrei Shleifer, Jeremy Stein, Jim Stock, Sam Thompson, and Tuomo Vuolteenaho for helpful comments, and Owen Lamont, Inmoo Lee, and Nejat Seyhun for data. Baker gratefully acknowledges financial support from the Division of Research of the Harvard Business School.

I. Introduction

Market timing is the tendency of firms to issue equity before low equity market returns. In contrast, *pseudo* market timing, as recently defined by Schultz (2003), is the tendency of firms to issue equity *following high* returns. In small samples, pseudo market timing can give the appearance of genuine timing. Consider an extreme example of pure pseudo market timing with only two returns. If the first return is high, equity issues rise; if the first return is low, equity issues fall. The first return can be mechanically explained *ex post*: Relatively low equity issues precede a high first return and relatively high equity issues precede a low first return. Even though the returns are random, equity issues "predict" returns in sample more often than not.

In a provocative article, Butler, Grullon, and Weston (2004a) argue that an aggregate version of the pseudo market timing bias explains why the variable in Baker and Wurgler (2000), the equity share in new equity and debt issues, predicts stock market returns in sample. While Butler et al. focus their critique on a particular link between financing patterns and stock returns, their general argument—that the pseudo market timing bias extends to time-series predictive regressions—is of considerably broader interest, because virtually every aggregate managerial decision variable that has been used in predictive regressions, not just equity issuance, is correlated with returns in the direction that induces a bias.

Consider the following examples. Seyhun (1992, 1998) and Lakonishok and Lee (2001) find that high aggregate insider buying appears to predict high stock market returns. However, aggregate insider buying increases as stock prices fall, raising the possibility that their result is driven by the bias. Baker, Greenwood, and Wurgler (2003) find that a high ratio of long-term debt issuance to total debt issuance appears to predict lower excess bond returns. However, as Butler et al. (2004b) observe, long-term debt issues are relatively more common as interest rates fall (and excess bond returns rise), again raising the potential for bias. Lamont (2000) finds that corporate investment plans forecast lower stock market returns, but planned investment increases with stock prices. And, finally, Baker and Wurgler (2004) find that the aggregate rate of dividend initiation is inversely related to the future returns on dividend payers over nonpayers. But again, the initiation rate increases with the relative returns on payers. Thus, in each of these papers, in addition to those involving equity issuance, the bias is a potentially serious concern. The size of the bias needs to be empirically pinned down before any of the above results can be accepted, or rejected, with confidence.[1]

In this paper, we empirically estimate the aggregate pseudo market timing bias that affects predictive regressions based on managerial decision variables. We start by observing that none of this is a fundamentally new question in asset pricing or time series econometrics. While Butler et al. (2004a) do not make the connection, aggregate pseudo market timing is simply a new name for the small-sample bias studied by Stambaugh (1986, 1999), Mankiw and Shapiro (1986), Nelson and Kim (1993), Elliott and Stock (1994), Kothari and Shanken (1997), Campbell and Yogo (2003), Amihud and Hurvich (2004), Lewellen (2004), Polk, Thompson, and Vuolteenaho (2004) and others. These studies focus on valuation ratios, like dividend yield or market-to-book, which exhibit an extreme and mechanical form of pseudo market timing—for example, when the market crashes, the dividend yield automatically rises. Our predictors are different, but the bias is the same, and it can be estimated using the same standard methods.

---

[1] It is important to note that the aggregate pseudo market timing bias discussed in this paper, and in Butler et al. (2004a,b), is a purely time-series phenomenon. It is not the bias emphasized by Schultz (2003), who discusses the potential for bias in "event time" studies of abnormal IPO returns. There, the problem arises when the *number* of firms going public increases following high abnormal returns on previous IPOs. For studies of that conceptually distinct type of bias, see Schultz (2003, 2004), Ang, Gu, and Hochberg (2004), Dahlquist and de Jong (2004), and Viswanathan and Wei (2004). For a general discussion, see Ritter (2003).

Specifically, when the predictor variable is stationary, and all of the managerial decision predictors we consider are theoretically stationary by construction, it is straightforward to run simulations and assess the magnitude of the small-sample bias induced by aggregate pseudo market timing. In these simulations, we impose the null hypothesis of no genuine market timing and varying degrees of pseudo market timing, mechanically tying the equity share and other candidate predictor variables to contemporaneous returns.

Our simulation results are, in a sense, a big letdown. Contrary to the conclusions of Butler et al. (2004a,b), the aggregate pseudo market timing bias is only a minor consideration for every variable we consider. The results can be described in terms of the theoretical determinants of the bias. As shown in Stambaugh (1986, 1999), the bias is most severe when the sample is small (on the order of $1/T$), the predictor is persistent, and its innovations are highly correlated with returns. As it turns out, empirically relevant values for these parameters are unable to generate a significant bias.

For example, aggregate pseudo market timing, of the degree observed in sample, has less than a one percent chance of reproducing the predictive power of the equity share variable. A reasonable point estimate is that one percent of its OLS coefficient is due to pseudo market timing. Even when we impose, counterfactually, pure pseudo market timing, forcing the correlation of innovations in the equity share and returns to one, approximately 90% of the OLS coefficient remains unexplained. And when we further increase the autocorrelation of the equity share by three standard deviations from its actual level, over 80% of the OLS coefficient still remains unexplained, and pseudo market timing of this sort still has less than a one percent chance of equaling the actual predictive coefficient. The bottom line is that in a sample of 75 years, small-sample bias is quite modest compared with the equity share's actual coefficient.

Results for other predictors are qualitatively similar—in no case we consider is the bias large enough to cast doubt on OLS-based inferences about predictive power. In summary, the aggregate pseudo market timing bias is a minor concern for the predictive regressions based on managerial decision variables that appear in the literature.

To reach closure on the issue, we briefly revisit the exercises in Butler et al. (2004a) that led them to conclude, apparently incorrectly, that small-sample bias drives the equity share's predictive coefficient. We note that Butler et al. do not present any direct estimates of the bias, as we do here, but instead build a case based on four indirect exercises. We observe that while some of these exercises represent robustness tests that may be of interest to an outside investor, others are problematic, and none is actually closely related to the small-sample bias issue. Thus, their results present no real conflict with the simulation results.

The paper proceeds as follows. Section II reviews the aggregate pseudo market timing bias and places it within an empirical framework that can be used to run simulations. Section III describes the data. Section IV reports simulation results. Section V comments further on the equity share variable. Section VI concludes.

II. Estimating the aggregate pseudo market timing bias

A. Empirical framework

An important conceptual point is that the aggregate pseudo market timing bias—in other words, the pseudo market timing bias in the context of time-series predictive regressions—is just a different name for an issue that is well understood in the financial econometrics literature. A common empirical framework is the system used by Mankiw and Shapiro (1986), Stambaugh (1986), and subsequent authors:

$$r_t = a + bX_{t-1} + u_t,\ u_t \sim \text{i.i.d. } (0, \sigma_u^2) \quad (1)$$

$$X_t = c + dX_{t-1} + v_t,\ v_t \sim \text{i.i.d. } (0, \sigma_v^2) \quad (2)$$

where r denotes returns on the stock market, for example, and X is a candidate predictor variable, such as aggregate equity issuance. Eq. (1) is the predictive regression, while Eq. (2) describes the evolution of the predictor. The contemporaneous covariance between the disturbances is σ_{uv}. We assume that the predictor is stationary, so that $|d|<1$.

To connect this framework to aggregate pseudo market timing, we adapt and (closely) paraphrase the following discussion from Stambaugh (1999, p. 379), in which he illustrates why the OLS estimate $\hat{b}$ is biased in its simplest possible setting. Consider repeated samples of only two observations, (r_1, X_0) and (r_2, X_1), so that $\hat{b}$ in each sample is just the slope of the line connecting these points, $\hat{b} = \frac{r_2 - r_1}{X_1 - X_0}$. Suppose $b = 0$, meaning that managers do not have genuine market timing ability; $d \approx 1$, so that innovations to equity issuance are highly persistent; and $\sigma_{uv} > 0$, meaning that managers are pseudo market timers, increasing equity issuance as stock prices rise. Consider those samples in which the first return is relatively low, $(r_2 - r_1) > 0$, or $u_2 > u_1$ (since $b = 0$). On average here, u_1 is negative, and because $\sigma_{uv} > 0$, v_1 is also on average negative, and so $(X_1 - X_0)$ is on average negative. Thus, $\hat{b} < 0$ on average. Now consider samples in which the first return is relatively high. On average here, u_1 is positive, hence v_1 is also on average positive, and therefore $(X_1 - X_0)$ is on average positive. Again, $\hat{b} < 0$ on average. Thus, on average across all samples, one will see a negative relation between equity issuance and subsequent returns, even though no timing ability exists. (In settings where $\sigma_{uv} < 0$, then $\hat{b} > 0$ on average.) This is the aggregate pseudo market timing bias.

This two-period example highlights the three main determinants of the bias. First, as the pseudo market timing correlation σ_{uv} goes to zero, the bias disappears, because the signs of u_1 and v_1 are no longer connected. Second, as the persistence of the predictor d goes to zero, the bias shrinks, because then the sign of $(X_1 - X_0)$ is less tightly linked to the sign of v_1 and thus to the sign of $(r_2 - r_1)$. But even when $d = 0$, there is still some correlation and thus some bias. Third, as the number of observations T increases, the bias approaches zero, because (with $b = 0$) the scatter of points becomes a horizontal cloud of these two-point clusters.

In the T-period case, Stambaugh (1986, 1999) shows that the size of the bias in $\hat{b}$ when u and v are distributed normally in the system above is

$$E[\hat{b} - b] = \frac{\sigma_{uv}}{\sigma_v^2} E[\hat{d} - d]\,. \tag{3}$$

The downward small-sample bias in the OLS estimate of d is shown by Kendall (1954) to be approximately $-(1 + 3d)/T$. So, mentally substituting this into Eq. (3), one sees that the pseudo market timing correlation, the predictor's persistence, and the sample size remain the key determinants of bias in the T-period case.

B. Simulation procedure

As mentioned in the Introduction, a large literature has considered the bias in $\hat{b}$ when X is a scaled-price variable like the aggregate dividend yield or book-to-market. Because dividends and book values are persistent, innovations in the dividend yield and the aggregate market-to-book ratio are highly correlated with contemporaneous returns, and thus an extreme, mechanical pseudo market timing correlation arises. While our predictors are different, the nature of the underlying bias is identical, and it can be estimated using the same empirical techniques developed in, for example, Nelson and Kim (1993) and Kothari and Shanken (1997).

In particular, when the predictor is stationary, it is straightforward to simulate Eqs. (1) and (2) to determine the size of the bias. The predictor variables we consider are theoretically stationary by construction (although in any given small sample, one might not be able to reject a unit root). In the simulations, we impose the null hypothesis of no predictability ($b = 0$) and vary the pseudo market timing correlation, σ_{uv}, and the other key parameters, d and T, to see whether a significant bias in $\hat{b}$ can be generated for empirically relevant parameters.

An example illustrates the basic procedure. In our benchmark simulations, we use the empirically most relevant parameter set: the bias-adjusted estimate of d ($\hat{d} + \frac{1+3\hat{d}}{T}$); the empirical distribution, and hence the correlation, of OLS estimates of u and v; and the number of observations actually available for the given predictor as T. We then simulate $100 + T$ values for r and X, starting with the actual X_0 and drawing with replacement from the empirical joint distribution of u and v. We throw away the first 100 values, leaving a sample size of T, which we use to compute a simulated OLS estimate $\hat{b}$. (When we examine three-year cumulative returns, the sample size for the regressions is $\frac{T}{3}$.) We repeat this procedure 50,000 times to plot the distribution of simulated OLS estimates, and then locate the actual estimate in this distribution. We then vary one or more of the parameters, generate a new simulated distribution, locate the actual OLS estimate there, and so forth.

An alternative approach is to compute reduced-bias p-values directly with the recently developed methods of Amihud and Hurvich (2004) and Polk, Thompson, and Vuolteenaho (2004). These two procedures lead to virtually identical inferences, and so we focus primarily on the simulation results, which allow us to consider situations where the degree of pseudo-market timing and the level of persistence in equity issues are counterfactually high.

III. Data

A. Predictor variables

We focus on six aggregate managerial decision variables. Five have previously been examined in a predictive regression context, and all six, based on the *a priori* considerations outlined in the Introduction, are likely to be subject to at least some degree of aggregate pseudo market timing bias. We replicate the methodology of the original studies where possible. When faced with a choice about the return prediction horizon, however, theory is no guide, so we use the horizons that are "strongest" or most emphasized in the original studies.

The first four predictor variables are used to forecast one-year-ahead real stock market returns. The equity share in new issues, i.e., the ratio of aggregate gross equity issuance to aggregate gross equity plus debt issuance, is derived from the *Federal Reserve Bulletin* data and is discussed in Baker and Wurgler (2000) and Butler et al. (2004a). Henderson, Jegadeesh, and Weisbach (2004) study the equity share variable using international data. By construction, this variable isolates the security choice decision from the level-of-external-finance decision. In the *Bulletin* data, equity issues include common and preferred, and debt issues include public and private. The annual series covers 1927 through 2001.

Detrended equity issuance is also based on the *Bulletin* annual gross equity issuance series. We take the log difference of aggregate gross equity issuance in year *t* and the average gross equity issuance over the previous five years (*t*-1 through *t*-6). This annual series covers 1932 though 2001. We are not aware of a prior study that uses exactly this variable.[2] It gives a different perspective on aggregate equity issuance than the equity share variable and will prove useful in our discussion of that variable.

---

[2] See Lamont (2002) on the predictive power of net new lists and Dichev (2004) on net equity capital flows.

Planned investment growth is studied by Lamont (2000). This series is based on a Commerce Department survey of firms' planned capital expenditure in the coming year. Lamont defines real planned investment growth as planned capital investment in year *t* divided by actual capital investment in *t*-1 all minus the growth in the national income accounts' nonresidential fixed investment deflator. As Lamont notes, investment plans for year *t* are reported as of February of *t*. The annual series of real planned investment growth for 1948 through 1992 is from his website.[3]

Aggregate insider buying is studied by Seyhun (1992, 1998) and Lakonishok and Lee (2001). Seyhun shared with us the monthly series, derived from the SEC's *Ownership Reporting System* file, on the fraction of publicly traded firms with net insider buying, as plotted in Seyhun (1998, p. 117). We average this series across months to construct an annual insider buying series from 1975 through 1994.

The long-term share in debt issues, i.e., the ratio of aggregate long-term debt issuance to aggregate short- plus long-term debt issuance, is used to forecast cumulative three-year excess returns on long-term government bonds in Baker, Greenwood, and Wurgler (2003). The debt issuance data is from the Federal Reserve *Flow of Funds*.[4] Short-term debt is primarily bank loans, and issuance is defined as the level of short-term credit market debt outstanding. Long-term debt is primarily corporate bonds. Under the assumption that one-tenth of long-term debt outstanding matures each year, issuance is defined as the annual change in the level of long-term debt outstanding plus one-tenth the lagged level. The long-term share series is annual from 1953 through 2000.

---

[3] http://www.som.yale.edu/faculty/oal4/research/publicplans.xls.

[4] This data comes from corporate balance sheets, includes a broader range of liabilities, and so does not match the totals from the *Federal Reserve Bulletin* that we use in computing the equity share in new issues.

The aggregate rate of dividend initiation, i.e. the percentage of the previous year's surviving nonpayers that paid positive dividends this year, is used by Baker and Wurgler (2004) to forecast three-year cumulative excess returns of dividend payers over nonpayers. They derive aggregate dividend payment series from aggregations of Compustat data. The dividend initiation series is annual from 1963 through 2000.

We refer the reader to the corresponding papers for more details on these variables. We omit summary statistics, because all of our analysis uses a standardized version of each predictor that has zero mean and unit variance across the full sample period.

B. Returns

Real annual stock market returns are based on the CRSP NYSE/Amex/Nasdaq value-weighted and equal-weighted return series, converted to real terms using the Consumer Price Index from Ibbotson Associates. The mean value-weighted (equal-weighted) real stock market return from 1928 through 2002 is 8.10% (13.30%) and the standard deviation is 20.47% (30.91%). Lamont's (2000) planned investment variable is dated the end of February, so we match it to subsequent March-February stock market returns.

Excess returns on long-term Treasury bonds over bills are from Ibbotson Associates; their government bond returns series uses data from the *Wall Street Journal* for 1977 through 2000 and the CRSP Government Bond File for 1976 and earlier. The mean three-year cumulative excess return for years starting with 1954 through 1998 (and ending with 1956 through 2000) is 2.06% with a standard deviation of 16.71%.

Finally, excess returns on dividend payers over nonpayers are based on the book-value-weighted return indexes of payers and nonpayers derived from CRSP and Compustat and described in Baker and Wurgler (2004). The mean three-year cumulative excess return on payers over nonpayers for years starting with 1964 through 1998 (and ending with 1966 through 2000) is 3.09% with a standard deviation of 37.99%.

We use r_t to denote the return in year t and R_{t+k} to denote the k-year cumulative return that starts with the year t return.[5]

IV. Simulation results

A. Predictive regressions based on equity issuance

Table 1 shows simulation results for predictive regressions using aggregate equity issuance variables for one-year-ahead stock market returns. Panel A reports simulations for the equity share and Panel B considers detrended equity issuance. The left side of the table contains the simulation inputs, which are always based on the null of no predictability (b = 0), and the right side reports the simulated distribution of the predictive regression coefficient and locates the actual OLS coefficient in this distribution. Figure 1 presents some of the interesting cases graphically. The basic approach is to start with empirically relevant parameter values, to determine the size of the bias in practice, and then proceed to progressively more extreme counterfactual parameter values to get a sense of what would be required for pseudo market timing to explain a given predictive coefficient.

Accordingly, in the first row of Panel A, we start with a benchmark parameter set that includes a bias-adjusted estimate of the equity share's autocorrelation, the empirical correlation of u and v (under b = 0), and a sample size of 75. The simulations indicate that the aggregate pseudo market timing bias is small, if not negligible, for this empirically relevant parameter set.

[5] The excess returns on long-term bonds (R_{GLt+3} - R_{GSt+3}) and dividend payers ($R_{Dt+3} - R_{NDt+3}$) are measured in logs and summed across overlapping three-year periods starting with year t:

$$R_{At+3} - R_{Bt+3} = \sum_{s=1}^{3} \log\left(\frac{1+r_{At+s}}{1+r_{Bt+s}}\right).$$

The actual OLS coefficient is -6.44 (and the unreported OLS heteroskedasticity robust t = -3.33). In contrast, the average simulated coefficient, under the null of no predictability, is only -0.07. Thus, as a point estimate, the bias accounts for 1.03% of the equity share's actual coefficient on value-weighted returns. Using the analytic expression for the bias in Eq. (3) leads to similar results, with $b^{bias\text{-}adjusted}$ = -6.34.[6] The one-tail p-value shows that there is only a 0.4% probability that the bias would lead to a coefficient as negative as the actual coefficient.[7] These results are presented graphically in Panel A of Figure 1. The actual coefficient, marked with a diamond, falls in the left tail of the simulated distribution.

In the second set of parameter values, we consider the counterfactual case of *pure* pseudo market timing, setting the correlation of u and v to one. To do this, we create a new v that is equal to u but standardized so that it has the same standard deviation as the empirical v series (i.e., $v_t = u_t \cdot \frac{s_v}{s_u}$ where s denotes sample standard deviation). The second row of Table 1 shows that even pure pseudo market timing generates only a small bias given empirical values of d and T. In particular, the mean simulated coefficient is only 11.37% as large as the actual coefficient, and there is only a 1.1% probability that extreme pseudo market timing of this type would lead to a coefficient as low as the actual coefficient. This case is plotted in Panel B of Figure 1.

---

[6] Amihud and Hurvich (2004) propose an estimator for the standard error of the bias-adjusted estimate. Because the distribution of b is not always symmetric, we present one- and two-tailed p-values from the simulations instead. None of the inferences in Tables 1 and 2 is materially affected. For example, the Amihud and Hurvich one-tailed p-values are 0.0036 for the case of value-weighted returns on the equity share in new issues and 0.0004 for the corresponding case with equal-weighted returns.

[7] Note that the simulations presume knowledge of the parameters d and σ_{uv}. We input the bias-adjusted estimate of d and use the empirical distribution of the OLS residuals u and v. So, the p-values in Table 1 and 2 are not exact. Polk, Thompson, and Vuolteenaho (2004) use a neural network to implement the theoretical result in Jansson and Moreira (2003), developing a conditional t-statistic and a set of critical values that do not require knowledge of d and σ_{uv}. Like the Amihud and Hurvich standard errors, this procedure produces almost identical inferences. For example, one-tailed p-values are 0.0044 for the case of value-weighted returns on the equity share in new issues and 0.0004 for the corresponding case with equal-weighted returns.

In the third parameter set, we again assume pure market timing, and try to further increase the bias by increasing d to three standard errors above its bias-adjusted estimate to 0.819. Even when one tilts these two parameters as far as possible toward pseudo market timing, the average simulated coefficient is still only 16.43% of the actual coefficient in a sample of 75 observations. The simulated distribution is in Panel C of Figure 1.

In the fourth and most extreme parameter set, we simultaneously consider pure market timing, a counterfactually high predictor autocorrelation, and a counterfactually low ten observations. Only in this extreme case is the aggregate pseudo market timing bias on the order of the actual coefficient. The simulated distribution is plotted in Panel D of Figure 1. The actual coefficient falls near the middle. This case highlights the fact that the bias is fundamentally a (very) small-sample problem.

The remaining rows in Panel A of Table 1 repeat these exercises for the case of equal-weighted market returns. The equity share has an actual coefficient of -11.92 (unreported OLS t = -3.49 and $b^{bias\text{-}adjusted}$ = -11.78) for equal-weighted returns. We modify the parameters as before, starting with the empirically relevant case and proceeding to extreme values. The biases here are also very small; under empirically relevant parameters, the point estimate is that the bias accounts for 0.83% of the actual coefficient. Panel B considers another measure of aggregate equity issuance, log deviation from a five-year moving average, which has an actual predictive coefficient of -7.83 (unreported OLS t = -4.05 and $b^{bias\text{-}adjusted}$ = -7.61) for value-weighted returns and -16.50 (unreported OLS t = -3.79 and $b^{bias\text{-}adjusted}$ = -16.19) for equal-weighted returns. The pattern of results is similar for this variable, indicating that it is also unaffected by the bias. For brevity, we do not present these results graphically.

In summary, the aggregate pseudo market timing bias has only a very small effect on predictive regressions based on aggregate equity issuance variables. The results clearly reject the conclusion of Butler et al. (2004a) that the equity share's actual predictive coefficient is due to pseudo market timing. In fact, only an extremely counterfactual parameter set can generate a bias that approaches the observed coefficient.

B. Predictive regressions based on other managerial decisions

Table 2 considers the effect of the aggregate pseudo market timing bias on a range of other predictors derived from managerial decisions. As suggested in the Introduction, many such predictors are likely to have, or already known to have, the correlation structure that induces a bias. It is an empirical question whether the bias is big enough to affect previous inferences. We report simulations using value-weighted returns in the first two panels of Table 2; results for equal-weighted returns are similar.

Panel A considers the real planned investment growth variable examined in Lamont (2000). The actual OLS coefficient for standardized real planned investment growth is -7.15 (unreported OLS t = -3.74 and $b^{bias\text{-}adjusted}$ = -7.11). We vary the parameter set as usual, starting with the empirically relevant case. The correlation between innovations in investment growth is positive but not that high, and the persistence of this predictor is low. The upshot is no bias. Indeed, even in the fourth row, where we consider the extreme counterfactual parameter set, there is only a 16.2% chance of observing a coefficient as negative as the actual coefficient under the null of no predictability.

Panel B looks at aggregate insider buying, as studied by Seyhun (1992, 1998) and Lakonishok and Lee (2001).[8] The actual predictive coefficient on standardized insider buying is

---

[8] Seyhun (1998) does not report formal predictive regressions. Instead, he reports a suggestive difference in average one-year-ahead returns conditional on aggregate "buy" signals (defined as 55% or more firms being net buyers over

4.99 (unreported OLS $t = 2.14$ and $b^{bias\text{-}adjusted} = 4.04$). Insiders are contrarian, as reflected in the highly negative correlation between insider buying and contemporaneous returns. Further, there are only 19 observations on this variable, further increasing the potential for small-sample bias. Indeed, although the predictor's autocorrelation is not high, this is the variable where the bias has the most bite. Nonetheless, the first row of Panel B reports the bottom line, which is that under empirically relevant parameters, the simulated coefficient equals or exceeds the actual coefficient only 10.5% of the time. As before, it takes extreme parameters to generate a bias as large as the actual coefficient.[9]

Panel C considers the long-term share in debt issues as a predictor of three-year excess returns on government bonds. Baker et al. (2003) find that when interest rates fall (and thus excess bond returns are high), corporations are issuing more long-term debt relative to short-term debt. But despite the strong autocorrelation in the long-term share, and only 15 non-overlapping observations, the bias turns out to be negligible.[10] In the benchmark parameter set, the average simulated coefficient is only 8.33% of the actual coefficient (similarly, the unreported $b^{bias\text{-}adjusted} = -9.27$). The simulated coefficient is as negative as the OLS coefficient only 2.9% of the time. Only extreme counterfactual parameters generate an average simulated coefficient in the ballpark of the actual coefficient. These results reject the conclusion of Butler et al. (2004b), who argue that the long-term share's coefficient is due to aggregate pseudo market timing.

---

the prior 12 months) and aggregate "sell" signals (defined as 45% or fewer firms being net buyers). Lakonishok and Lee (2001) do present formal tests (for example, p. 96).

[9] Increasing the autocorrelation of insider buying by three standard errors above the bias-adjusted d leads to a value near or above 1.00. The same happens with the long-term share. Such an autocorrelation can be ruled out on theoretical grounds, so we cap the counterfactually high d value at 0.95. This is still well over two standard deviations above the bias-adjusted OLS estimate of d.

[10] We do not report the overstated OLS t-statistics from regressions with overlapping returns.

Finally, Panel D considers the aggregate rate of dividend initiation as a predictor of the three-year cumulative excess return on dividend payers over nonpayers. This is the regression emphasized in Baker and Wurgler (2004). The residual correlation indicates that when the return on payers is relatively high, nonpayers initiate dividends at a high rate, suggesting the possibility of bias. In addition, only 12 non-overlapping observations are available here. However, the simulations indicate that there is only a 4.4% chance of observing the actual coefficient under the empirical parameters. In the benchmark parameter set, the average simulated coefficient is only 9.48% of the actual coefficient (similarly, the unreported $b^{bias\text{-}adjusted}$ = -22.21). In this setting, it is difficult to generate a large bias even under extreme counterfactual parameters.

In summary, the aggregate pseudo market timing bias is only a minor concern for the predictive regressions based on managerial decision variables that appear in the literature. Of course, although we have found no case in which inferences based on OLS estimates would be seriously misleading, good practice dictates that the potential for small-sample bias be evaluated on a case-by-case basis going forward.

V. On the robustness of the equity share variable

The simulation results show rigorously that aggregate pseudo market timing has little effect on the equity share variable in particular, and therefore that the conclusion of Butler et al. (2004a) is incorrect. To draw the issue to a close, we briefly review how those authors reached their conclusion, and where they may have erred.

Butler et al. present no actual estimates of the bias. Indeed, they do not even mention that it is a small-sample problem, which eventually disappears regardless of the pseudo market timing correlation. Instead, they build a case based on four indirect exercises. We discuss these in turn.

We will show that while one or two of the Butler et al. exercises can be seen as useful robustness tests, of interest to an outside investor aiming to exploit the phenomenon, none presents a strong link between the equity share's predictive coefficient and small-sample bias, and so there is ultimately no conflict with our simulations.

A. Sensitivity to "outliers"

Butler et al. begin by presenting evidence that the equity share and returns are positively correlated around a dozen historical stock market events (see their Figures 1 through 3 and Table 4). As they point out, and as is confirmed more systematically in our Table 1, such a correlation structure suggests the possibility of pseudo market timing bias. Butler et al. then remove five years of the sample, 1929, 1930, 1931, 1973, and 1974, arguing that the events of these years, the Great Crash and the oil crises, were unpredictable and hence the associated fall in the equity share that occurs in these years could not be due to genuine market timing. Observing that the equity share is insignificant in the remaining sample (see their Table 3), the authors propose that aggregate pseudo market timing, centered on these five years, is driving the full-sample predictive relationship.

This exercise makes a statement about robustness, but sheds little light on small-sample bias. There are several reasons why this link cannot be made. First, a positive correlation between returns and the equity share is equally predicted by *both* the genuine and pseudo market timing hypotheses. Under pseudo market timing, the fall in the equity share around large negative returns is the definition of the effect; under genuine market timing, a fall in prices implies less overvaluation, and hence less of a reason to prefer equity. Thus, the simple fact that the equity share and returns both fall in a given period does not identify the observation as any more likely to be associated with pseudo market timing.

Second, although Butler et al. offer a list of events as having the potential to induce bias (see their Table 4), they only delete years corresponding to two such events. But this arbitrary choice turns out to be important. If one removes five *additional* years from the sample, 1940, 1941, 1987, 1990, and 2001, all of which appear on their own selected list of "unanticipated events" of concern, a significant coefficient actually reappears.[11] But in what sense are the events that drove stocks in these five years—the German invasion of France, Pearl Harbor, the Iraqi invasion of Kuwait, the Crash of 1987, and September 11th—any more or less predictable than the Great Crash and the oil crises?

Third, and most fundamentally, under the null hypothesis *no* return is predictable. The notion that predictability can be tested by throwing out individual data points based on a qualitative judgment that certain events are unpredictable seems strange from the start. After all, the aim is to *test* the null hypothesis of no predictability. Indeed, given the nature of the market timing hypothesis, it seems particularly strange to throw out potentially highly informative periods of possible market correction, like the Great Crash.

Nonetheless, even though these five data points cannot be attached to pseudo as opposed to genuine market timing, it does seem striking that removing them reduces the OLS coefficient by as much as 57.61%.[12] As a generic observation about robustness, this still seems damaging. Yet, as it turns out, even this observation is not surprising. This is close to the expected outcome from the strategic removal of five data points from the southeast quadrant of a 75-point

---

[11] The predictive regression coefficient is -4.86 (t = -1.67, $p<0.10$) for value-weighted returns and -11.62 (t = -2.09, $p<0.05$) for equal-weighted returns.

[12] For value-weighted returns, dropping the five data points attenuates the coefficient by 57.61%, from -6.44 to -2.73. For equal-weighted returns, the attenuation is 27.62%, from -11.91 to -8.62). Butler et al. report that the coefficient drops all the way to -2.06 for value-weighted returns and to -6.53 for equal-weighted returns (see their Table 3), but this is because they restandardize the equity share after dropping five data points. Because these were also extreme values of X, dropping them reduces the standard deviation, increases the variance of the restandardized variable, and lowers the coefficient. To allow for an apples-to-apples comparison, we do not restandardize.

scatterplot that begins with an R^2 of only 10% to 15%. To illustrate this formally, we run simulations to trace out the expected attenuation of a predictive regression coefficient under various "outlier" removal schemes.

The design choice in these simulations is how to characterize the five data points that Butler et al. actually remove. Once that is done, we can replicate the spirit of the process and estimate the associated attenuation. We consider two data-removal schemes. The first springs from the observation that the data points Butler et al. remove include five of the seven lowest value-weighted returns, and five of the six lowest equal-weighted returns. *Pseudo market timing or not*, it is well known that removing extreme values of the dependent variable (henceforth r to match our earlier notation) will attenuate the OLS coefficient.[13]

Accordingly, in our first set of simulations, we draw samples of 75 data points, where the independent variable (henceforth X) and r match the R^2 of the predictive regression of the equity share for value-weighted and equal-weighted market returns. We then remove observations with the most extreme realizations of r. Note that this scheme is somewhat more aggressive than the Butler et al. procedure in the sense that it picks off exactly the most extreme returns, but it is less aggressive in that it does not identify "outliers" based on X, only on r.

The second data removal scheme is more strategic, and more closely matches what Butler et al. have effectively done. As it turns out, they delete five data points from the southeast quadrant of the scatterplot. Thus, the data are identified from both realizations of X and r. This

[13] For example, see Goldberger (1981) for the derivation of a simple analytic solution for the OLS coefficient b in a truncated sample:

$$b^* = \frac{\theta}{1-\rho^2(1-\theta)} b$$

where θ is the ratio of the variance of r in the truncated sample to the variance of r in the full sample and ρ^2 is the coefficient of determination, or the population R^2. The attenuation is largest when the R^2 is low and the truncation leads to a large reduction in the variance of r. This characterizes the large-sample attenuation, but provides some intuition for the situations where we would expect to find bias in a small sample.

scheme is likely to attenuate a negative coefficient—again, pseudo market timing or not—even more than the first scheme.

In the second set of simulations, we start by removing the data point from the southeast quadrant (or equivalently from the northeast quadrant if X and r are positively correlated) with the most negative value of $-X_t \cdot e_t$. This is the most influential point in the southeast quadrant. In other words, it exerts the most downward pressure the regression line. Next, we refit the line with the remaining 74 data points, and remove the new most influential point from the southeast quadrant, the new most negative value of $-X_t \cdot e_t$. We continue this process of refitting the regression line and removing data. When applied to the original data set, the first five data points removed by this process are *precisely* the five years removed by Butler et al.[14] Therefore, this scheme should yield a fairly accurate approximation of the attenuation that is to be expected upon removing data points "like" 1929, 1930, 1931, 1973, and 1974.

Figure 2 shows the expected attenuation, and 95% confidence intervals, under these data removal schemes. Panels A and B show simulations set up to match the case of value-weighted returns, and Panels C and D consider equal-weighted returns. In each panel, the diamond shows the actual attenuation Butler et al. find upon removing 6.67% (5/75) data points. The basic finding is that the observed attenuation is unsurprising. For instance, Panel C shows that when the 6.67% most extreme returns are deleted from scatterplots corresponding to the case of equal-weighted returns, the expected attenuation is almost exactly that which Butler et al. actually find. In the case of value-weighted returns in Panel A, the diamond is also within the 95% confidence band. Panels B and D are more relevant, because they better capture Butler et al.'s effective data removal scheme. The expected attenuation grows even faster in this scheme. And for both equal-

[14] We do not suggest that this is how Butler et al. actually selected the five data points they remove. Our goal is simply to characterize, as accurately as possible, data points "like" those observations.

and value-weighted returns, the observed attenuation is not unusual. It is larger than expected for value-weighted returns, though easily within the confidence interval, and even a bit smaller than expected for equal-weighted returns.[15]

Thus, the first Butler et al. exercise does not tell us much about either small-sample bias or sensitivity to outliers. However, it does raise a legitimate question as to whether the equity share is driven by a small number of years. One approach to answering this question is to remove the most extreme values of X, as opposed to data points least consistent with the null hypothesis. The idea is that this process will mechanically increase standard errors by reducing the variance of the predictor, but, if the predictive power is widespread, should not attenuate the regression coefficient. Using this approach, we find that removing the top and bottom equity share years changes the coefficient in the value-weighted predictive regression from -6.44 to -5.88; removing the top and bottom five years gives a coefficient of -3.52; and even removing the top and bottom ten years gives a coefficient of -4.94. The coefficient is negative for every case in between. Figure 3 shows similar results for the equal-weighted case: There, the coefficient is actually larger once one removes the 20 most extreme realizations of the equity share. Clearly, this exercise shows that the coefficient is not sensitive to a few extreme equity share years.

Another standard approach to examining the effect of extreme values is to use median regression. Coefficient estimates in median regressions are both larger in absolute value, at -7.51 (t = -2.39) and -13.20 (t = -3.88), respectively, than they are in OLS regressions.

We have made several points in this section. To summarize, the observations that Butler et al. remove cannot logically be attached to pseudo market timing as opposed to genuine market

---

[15] The expected values in Panels B and D are still somewhat understated. Butler et al. remove the five most influential data points from the southeast quadrant of the scatter plot, but these turn out to be five of the seven most influential value-weighted data points and five of the six most influential equal-weighted data points *overall*, when we consider observations from both the northwest and southeast quadrants. The simulation exercise only removes data from the southeast quadrant, which on average will only contain five of the ten most influential data points.

timing; the whole approach of identifying small-sample bias by removing data points that are more or less "predictable" is dubious; the sensitivity to outliers, as Butler et al. identify them, is not unusual for such low R^2 regressions; and, more standard robustness tests indicate that the predictive power of the equity share is fairly widespread.

B. Post-1975 period

In their second exercise, Butler et al. (2004a) suggest that the evidence for market timing should be stronger after the inclusion of Nasdaq in the CRSP aggregates, but in fact the equity share has little predictive power in this subsample (see their Table 5). Butler et al. offer this as indirect evidence that the pseudo market timing bias drives the full-sample relationship.

While this exercise again sheds no direct light on small-sample bias, it is a useful observation about robustness. One can learn more by probing further. Note that the equity share controls for total external finance by construction, but the cost of this virtue is that the variable is sensitive to debt market conditions. And, as it turns out, the post-1975 period has seen a secular shift toward debt issues. In the mid-1980s, debt issues increase almost fivefold, from 68 to 328 billion dollars, in the five years from 1983 to 1988 (see Baker and Wurgler (2000) Figure 1). Baker and Wurgler speculate that the growth in debt issues arises because of the emergence of the junk bond market and a shortening in average debt maturities. Neither has much to do with the state of the public equity markets, and so the equity share has become a less useful predictor in recent years as a result. For example, at the end of 1999, the equity share was at 0.14, well below its full-sample (or post-1975) mean, but few would characterize the equity issuance market as "below average" at this point.

The growth in equity issues from a five-year moving average is not contaminated by the secular shift in debt issues, and so is arguably a more consistent indicator of equity issuance activity over the last few decades. Table 3 shows that detrended equity issues predicts both equal- and value-weighted returns in the post-1975 period. In fact, the coefficient in the recent period is larger in absolute value—at -12.28 for value-weighted returns and -21.47 for equal-weighted returns—than for the full sample period. The *t*-statistics are smaller, at -1.93 and -1.99, in large part because the recent sample is only 28 years long. In summary, while the equity share has not been useful in the recent period, this appears to be due to developments in the debt market that contaminate the variable, not to a general breakdown in the relationship between equity issuance and future returns.

C. Separating debt and equity issues

Genuine market timing logically implies the substitution of debt, if it is more efficiently priced, for equity, Butler et al. argue. In their third exercise, they partition the sample into four parts according the growth rate of equity and debt issuance over the past year. There is little difference in the future stock market returns across the four groups. In particular, the difference in value-weighted returns between high equity growth/low debt growth years and low equity growth/high debt growth years is a statistically insignificant -4.14% (see their Table 6).

This is a peculiar test, because by definition the equity share already captures a first-order substitution effect. A high value of the variable indicates a high volume of equity issues relative to debt issues, and this in turn forecasts low equity returns, much as Butler et al. envision should occur. Their notion of substitution is akin to a second derivative. That is, they sort on the relative growth rates of equity and debt issues, which amounts to a sort on the change in the equity share—when equity issues grow faster than debt issues, the equity share goes up, and vice-versa. But the change in the equity share does not predict returns nearly as well as the level, and we can see no reason that it should.[16] In any case, for present purposes the main point is that this test again has little direct bearing on small-sample bias.

D. Out-of-sample forecasts

In their fourth and final analysis, Butler et al. consider the equity share's out-of-sample predictive ability. They point out that under pseudo market timing, no out-of-sample predictability should be present. They report that the equity share has little out-of-sample predictive power (see their Table 7), consistent with pseudo market timing.

The authors fail to mention that out-of-sample predictability is not a necessary implication of genuine market timing, either, so the test is again not very relevant. The question asked by Baker and Wurgler (2000) is whether managers systematically issue more equity before low market returns. For this hypothesis to be true, there need not be any out-of-sample predictability.

For example, suppose that an increase in the equity share genuinely predicts lower future returns in every subsample, but the *degree* of predictability varies over time. This would be the case if there were occasional and significant corrections. (It would also be the case if the aggregate market were more efficiently priced in some parts of the sample than in others.) Within a period that contains a correction, the predictability or, to be precise, the coefficient in a regression of future returns on the equity share will be large in absolute value. Within a period that does not contain a significant correction, the coefficient will be small, but still negative.

[16] To the extent that studying the relative growth rates of equity and debt issuance is informative, we make two further comments. First, the Butler et al. test gives more favorable results when one defines growth in issuance as a difference from a five-year moving average. This may give a better match with the frequency and duration of actual financing cycles. Using this definition, the difference between the off-diagonal groups becomes -9.89% for value-weighted returns and -19.50% for equal-weighted returns, and the latter difference is significant at the five percent level. Second, the growth rates of equity and debt issues are positively correlated, but a horse race predictive regression suggests that it is indeed growth of equity issues that predicts returns. Tables are available on request.

In this example, the power of the equity share to predict future returns is clearly genuine, and yet a test for out-of-sample predictability may find none. The reason is that the out-of-sample approach means using a *small* coefficient, from past 10-year rolling estimations for example, to explain periods where the level of predictability is *high*. The small coefficient is a bit better than a simple past average, but not much. More damaging to out-of-sample performance, the approach also means using a *large* coefficient to explain periods of returns where the level of predictability is *low*. This can be much worse than a simple past average. So, it is conceivable that equity issuance could reliably predict returns within each subsample, benefiting managers and existing shareholders, while at the same time the information contained in the equity share might not be nearly as valuable to outside investors.

The correct test of whether managers systematically issue more equity before low market returns is simply to look back and ask how managers performed. And by this metric, as reported in Baker and Wurgler (2000), they timed their equity issues fairly well when measured over the whole sample period. The bottom line is that the out-of-sample test, while likely of interest to an outside investor, does not even begin with the power to reject genuine market timing in favor of pseudo market timing.[17]

As far as the actual out-of-sample predictability of the equity share is concerned, it is worth a brief note that other authors have found considerably more than Butler et al. For instance, Goyal and Welch (2004) study fifteen annual time-series predictors, and report that the best (and only statistically significant) out-of-sample predictor is the equity share. Rapach and Wohar (2004) study nine annual time-series predictors, and again report that for one-year-ahead stock returns, the strongest out-of-sample performance is turned in by the equity share.

[17] Goyal and Welch (2004) also make this observation. "Baker and Wurgler (2000) are interested in the behavior of firms themselves, not with the prediction that outside investors may follow. Thus, they are appropriately interested in in-sample outperformance, not out-of-sample outperformance." (p. 5)

An effort to reconcile these claims is beyond the scope of this paper, because the outcome would shed little light on small-sample bias anyway, but Table 4 reports a simple test. We use all historical data available in each year, starting with 1940, to form out-of-sample predictions with the equity share and with a simple past average. Panels A and B show that the equity share reduces the annual root mean squared prediction error by 0.40% for value-weighted returns and 0.89% for equal-weighted returns.[18] The in-sample reduction in root mean squared error in the post-1940 period is only slightly larger, at 0.46% and 1.20%, respectively. Note that the only area where the equity share performs poorly out of sample is the post-1975 period, which is exactly where it performs poorly in sample. (Note that an in sample regression cannot increase the root mean squared error.) Panels C and D shows that when using the growth in gross equity issues from a five-year moving average, the out of sample performance even in recent decades is respectable. This indicates that the equity share, and equity issues more generally, perform similarly well out of sample and in sample.

E. Summing up

We have considered each line of analysis in Butler et al. (2004a), the study that motivated our investigation of aggregate pseudo market timing. We find their selective pruning of outliers problematic on several grounds. Their post-1975 results point to a decline in the usefulness equity share variable, but not to a breakdown in the broader relationship between aggregate equity issuance and returns. Their test for a substitution effect overlooks the fact that the equity share already captures a more basic notion of substitution, and their out-of-sample exercise does not even start with power to reject genuine market timing in favor of pseudo market timing.

[18] This is opposite to the results of Butler et al., who tend to find better results for the unconditional (constant) model. The difference appears to reflect their choice to use a rolling pre-estimation window, as opposed to all of the available historical data. Because the sample begins in 1927, this design choice has the effect of discarding the information from the earliest observations, including the Great Crash, from most of the analysis.

Ultimately, because these exercises are not closely related to the bias of interest, they do not present a conflict with our simulation results. All of the evidence that actually speaks to the question indicates that the predictive power of aggregate equity issues for stock market returns is not due to small-sample bias.

VI. Conclusion

The aggregate version of Schultz's (2003) pseudo market timing bias represents a potentially serious concern for many studies that use aggregate managerial decision variables to predict market-level returns. Butler et al. (2004a) consider this bias in the specific context of the equity share in new issues variable, but there are also a number of other results in the prior literature that *a priori* are equally likely to be affected by the bias. These include predictive regressions based on planned investment growth, the maturity of new debt issues, aggregate insider buying, and the aggregate dividend initiation rate.

In this paper, we start by pointing out that the aggregate pseudo market timing bias is a new name for the small-sample bias long known to affect predictive regressions based on scaled-price predictors such as the dividend yield. Prior work shows how to estimate the bias using simulation methods. We run such simulations, and find that the bias is, in practice, minor for all of the predictive regression settings we consider. Indeed, small-sample bias is particularly negligible for the equity share variable. Overall, the results bolster a growing body of evidence that managerial decisions respond to securities market conditions.[19]

[19] To what extent successful market timing is intentional as opposed to a more passive response to investor demand or market liquidity is an open question. For an example, see Baker and Stein (2004).

References

Amihud, Yakov, and Clifford Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis*, forthcoming.

Ang, Andrew, Li Gu, and Yael V. Hochberg, 2004, Is IPO underperformance a peso problem?, Columbia University working paper.

Baker, Malcolm, Robin Greenwood, and Jeffrey Wurgler, 2003, The maturity of debt issues and predictable variation in bond returns, *Journal of Financial Economics* 70, 261-291.

Baker, Malcolm, and Jeremy Stein, 2004, Market liquidity as a sentiment indicator, *Journal of Financial Markets* 7, 271-299.

Baker, Malcolm and Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns, *Journal of Finance* 55, 2219-2257.

Baker, Malcolm, and Jeffrey Wurgler, 2004, A catering theory of dividends, *Journal of Finance* 59, 271-288.

Butler, Alexander W., Gustavo Grullon, and James P. Weston, 2004a, Can managers forecast aggregate market returns?, *Journal of Finance*, forthcoming.

Butler, Alexander W., Gustavo Grullon, and James P. Weston, 2004b, Can managers successfully time the maturity structure of their debt issues?, Rice University working paper.

Campbell, John Y., and Motohiro Yogo, 2003, Efficient tests of stock return predictability, Harvard University working paper.

Dahlquist, Magnus, and Frank de Jong, 2004, Pseudo market timing: Fact or fiction?, University of Amsterdam working paper.

Dichev, Ilia D., 2004, What are stock investors' actual historical returns?, University of Michigan working paper.

Elliott, Graham, and James H. Stock, 1994, Inference in time series regression when the order of integration of a regressor is unknown, *Econometric Theory* 10, 672-700.

Goldberger, Arthur, 1981, Linear regression after selection, *Journal of Econometrics* 15, 357-366.

Goyal, Amit, and Ivo Welch, 2004, A comprehensive look at the empirical performance of equity premium prediction, Emory University working paper.

Henderson, Brian J., Narasimhan Jegadeesh, and Michael S. Weisbach, 2004, World markets for raising new capital, University of Illinois working paper.

Jansson, Michael, and Marcelo Moreira, 2003, "Conditional Inference in Models with Nearly Nonstationary Regressors," Harvard University working paper.

Kendall, M. G., 1954, Note on bias in the estimation of auto-correlation, *Biometrika* 41, 403-404.

Kothari, S. P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time-series analysis, *Journal of Financial Economics* 44, 169-203.

Lakonishok, Josef, and Inmoo Lee, 2001, Are insider trades informative?, *Review of Financial Studies* 14, 79-111.

Lamont, Owen, 2000, Investment plans and stock returns, *Journal of Finance* 55, 2719-2745.

Lamont, Owen, 2002, Evaluating value-weighting: Corporate events and market timing, Yale University working paper.

Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics*, forthcoming.

Mankiw, N. Gregory, and Matthew Shapiro, 1986, Do we reject too often? Small sample properties of tests of rational expectations models, *Economics Letters* 20, 139-145.

Marriott, F.H.C., and J.A. Pope, 1954, Bias in the estimation of autocorrelations, *Biometrika* 41, 390-402.

Nelson, Charles and M. Kim, 1993, Predictable stock returns: The role of small-ample bias. *Journal of Finance* 48, 641-661.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho, 2004, "Cross-sectional Forecasts of the Equity Premium," Harvard University working paper.

Ritter, Jay, 2003, Investment Banking and Securities Issuance, in Constantinides, George, Milton Harris, and René Stulz (eds.), *Handbook of the Economics of Finance* (Amsterdam: Elsevier North-Holland).

Rapach, David E., and Mark E. Wohar, 2004, In-sample vs. out-of-sample tests of stock return predictability in the context of data mining, University of Nebraska working paper.

Schultz, Paul, 2003, Pseudo market timing and the long-run underperformance of IPOs, *Journal of Finance* 58, p. 483-517.

Schultz, Paul, 2004, Pseudo market timing and the stationarity of the event-generating process, University of Notre Dame working paper.

Seyhun, H. Nejat, 1992, Why does aggregate insider trading predict future stock returns? *Quarterly Journal of Economics* 107, 1303-31.

Seyhun, H. Nejat, 1998, *Investment Intelligence from Insider Trading*, (Cambridge, MA: MIT Press).

Stambaugh, Robert F., 1986, Bias in regressions with lagged stochastic regressors, University of Chicago working paper.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

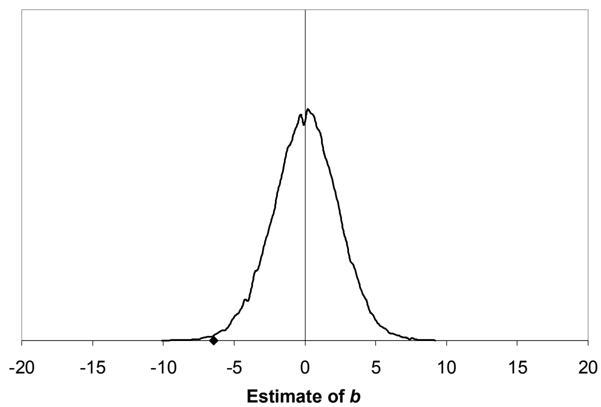
Viswanathan, S., and Bin Wei, 2004, Endogenous events and long run returns, Duke University working paper.

Figure 1. Pseudo-market timing and predictability. We simulate 50,000 estimates from the following system of equations under the null hypothesis of no predictability (b=0):

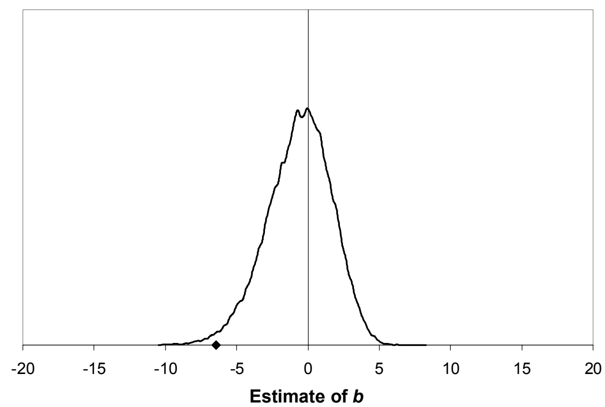
$$r_{mt} = a + bE_{t-1} + u_t,\ E_t = c + dE_{t-1} + v_t$$

where r is the value-weighted stock market return and E is the ratio of equity issues to total equity and debt issues. See Table 1 for details. The graph is a histogram of realizations of b. The diamond shows the location of the actual OLS estimate on the X axis. We impose the null hypothesis of no predictability (b = 0) in all cases. Otherwise, the parameters vary as follows. Panel A uses the in-sample bias-adjusted OLS estimate for d and the OLS estimate for ρ_{uv}. Panel B uses the in-sample OLS estimate for d and sets ρ_{uv} equal to one, leaving the standard deviation of u and v unchanged. Panel C increases the bias-adjusted OLS estimate for d by three OLS standard deviations. Panel D reduces the sample size to 10.

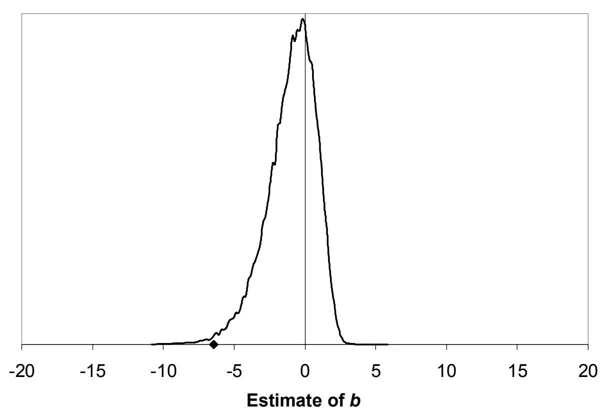
Panel A. d=0.507, ρ_{uv}=0.138, T=75



Panel B. d=0.507, ρ_{uv}=1.000, T=75



Panel C. d=0.819, ρ_{uv}=1.000, T=75



Panel D. d=0.819, ρ_{uv}=1.000, T=10

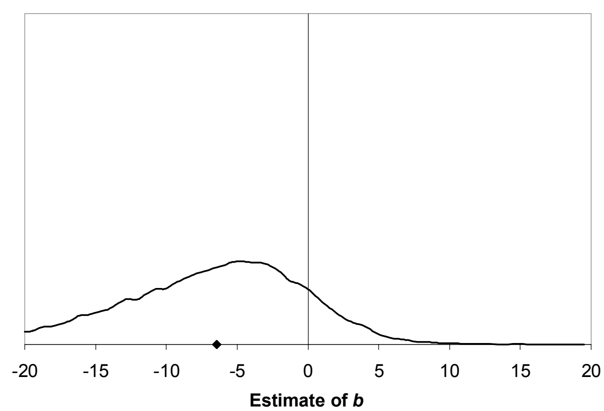
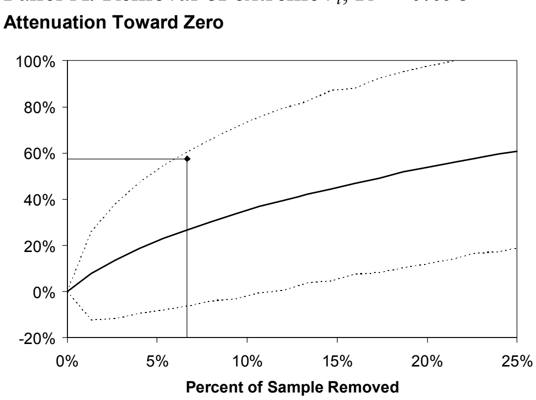


Figure 2. Removing outliers from low R-squared regressions. We simulate estimates of a univariate regression with 75 observations where the independent variable and the errors are drawn from normal distributions:

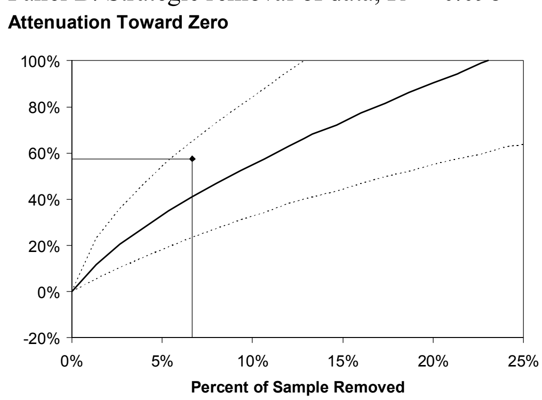
$$r_t = a + bX_t + e_t \text{, where } X \sim N(0, \sigma_X) \text{ and } e \sim N(0, \sigma_e) \text{, and where } N = 75$$

We choose relative values of $b<0$, σ_X and σ_e to match the R^2 of the regressions of aggregate value-weighted returns ($R^2 = 0.098$ in Panels A and B) and equal-weighted market returns ($R^2 = 0.149$ in Panels C and D) on the ratio of equity issues to total equity and debt issues. We draw 100,000 sets of 75 observations, throwing out sets with R^2's that fall outside a narrow band ($0.094<R^2<0.102$ for Panels A and B; $0.144<R^2<0.154$ for Panels C and D). Bands are constructed to contain 10% of R^2's, so roughly 10,000 sets of 75 observations remain. The solid line shows the mean and the dashed lines show a 95 percent confidence interval for the attenuation toward zero of the OLS estimate of b. In Panels A and C, we remove extreme values of r_t, starting with the largest in absolute value. In Panels B and D, we strategically remove extreme values of r_t, starting with the largest realization of $-X_t \cdot e_t$ where X_t is positive (the most influential point from the southeast quadrant of the r-X scatterplot) and sequentially removing the next largest realizations of $-X_t \cdot e_t$ from the sequentially re-estimated regressions. The diamond shows the combination of data removed and the reduction in the OLS estimate when Butler et al. (2004) remove five (or 6.67%) of 75 observations.

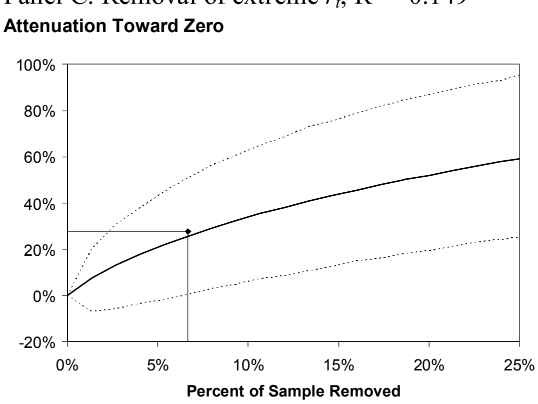
Panel A. Removal of extreme r_t, $R^2 = 0.098$



Panel B. Strategic removal of data, $R^2 = 0.098$



Panel C. Removal of extreme r_t, $R^2 = 0.149$



Panel D. Strategic removal of data, $R^2 = 0.149$

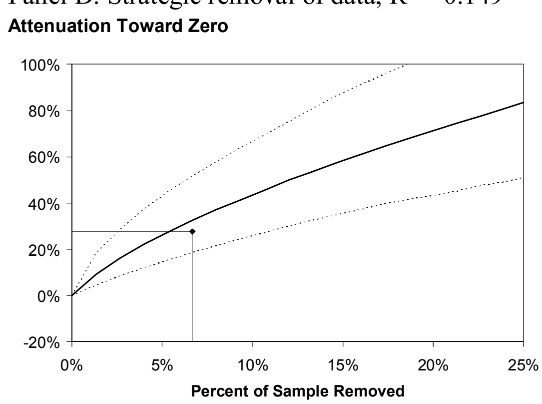


Figure 3. Removing extreme values of the equity share. We estimate the relationship between the equity share in new issues and future value-weighted and equal-weighted market returns, progressively removing the most extreme values of the equity share:

$$r_{mt} = a + bE_{t-1} + u_t .$$

The lines starts at the left with the full sample OLS coefficient (-6.44 for value-weighted returns, and -11.92 for equal-weighted returns). We remove the data points containing the top and bottom values of the equity share and re-estimate the OLS coefficient with the remaining sample; then we remove the top and bottom *two* data points; and so on. The solid line uses value-weighted returns and the dashed line uses equal weighted returns.

Value-weighted (solid) and equal-weighted (dashed) results

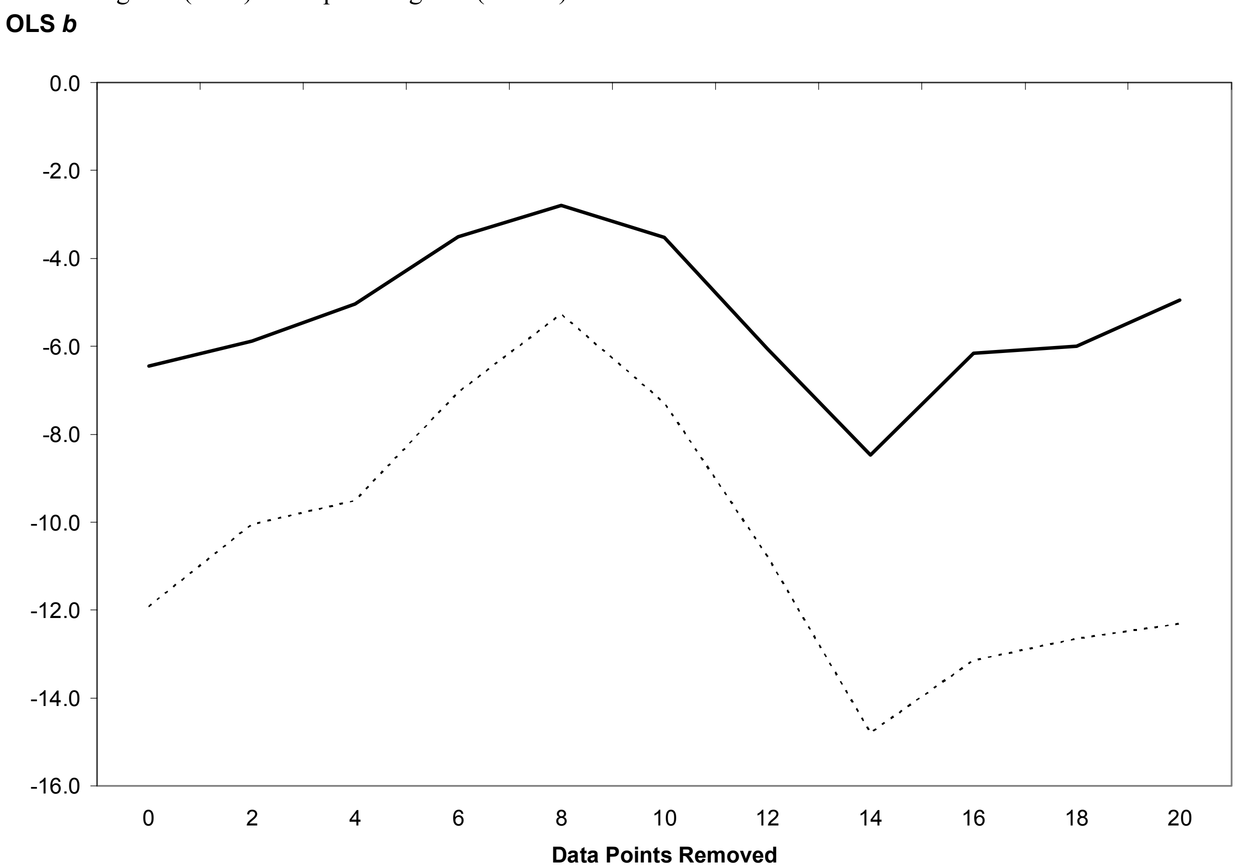


Table 1. Pseudo-market timing and predictability. We simulate 50,000 estimates from the following system of equations:

$$r_{mt} = a + bE_{t-1} + u_t \,,\; E_t = c + dE_{t-1} + v_t$$

where r is the aggregate market return, either equal- or value-weighted, and E is one of two measures of equity issues. The equity share in new issues is the ratio of equity issues to total equity and debt issues. Detrended equity issues is the log difference between the level of equity issues and the average level of equity issues in the previous five years. Both are standardized to have zero mean and unit variance. We simulate 100+T values for r and E starting with the actual E_0 and drawing with replacement from the empirical joint distribution of u and v. We throw away the first 100 values, leaving us with a sample size of T. We use OLS estimates of b for 50,000 separate samples, reporting the average and locating the actual OLS estimate in this simulated distribution with one- and two-tailed p-values. We impose the null hypothesis of no predictability (b = 0) in all cases. The parameters are as follows. The first row in each panel uses the in-sample bias-adjusted OLS estimate for d and the OLS estimate for ρ_{uv}. The second row uses the in-sample OLS estimate for d and sets ρ_{uv} equal to one, leaving the other distributional properties of u and v unchanged. The third row increases the bias-adjusted OLS estimate for d by three OLS standard deviations. The fourth row reduces the sample size to 10. The fifth through eighth rows repeat this for equal-weighted returns.

	Parameter Inputs			Simulation versus Actual Results				
	d	**ρ_{uv}**	**T**	**Average Coefficient**	**Actual Coefficient**	**Average/ Actual (%)**	**One-Tail p-value**	**Two-Tail p-value**
Panel A. Equity Share in New Issues = EI / (EI + DI)								
VW	0.507	0.138	75	-0.07	-6.44	1.03	[0.004]	[0.007]
VW	0.507	1.000	75	-0.73	-6.44	11.37	[0.011]	[0.011]
VW	0.819	1.000	75	-1.06	-6.44	16.43	[0.007]	[0.007]
VW	0.819	1.000	10	-7.38	-6.44	114.59	[0.511]	[0.519]
EW	0.507	0.132	75	-0.10	-11.92	0.83	[0.001]	[0.001]
EW	0.507	1.000	75	-1.08	-11.92	9.10	[0.002]	[0.002]
EW	0.819	1.000	75	-1.52	-11.92	12.79	[0.001]	[0.001]
EW	0.819	1.000	10	-10.73	-11.92	90.05	[0.406]	[0.411]
Panel B. Detrended Equity Issues = log (5·EI / ($EI_{-1}+EI_{-2}+EI_{-3}+EI_{-4}+EI_{-5}$))								
VW	0.581	0.201	70	-0.24	-7.83	3.10	[0.010]	[0.015]
VW	0.581	1.000	70	-1.16	-7.83	14.79	[0.020]	[0.020]
VW	0.801	1.000	70	-1.49	-7.83	18.99	[0.016]	[0.016]
VW	0.801	1.000	10	-9.80	-7.83	125.22	[0.543]	[0.554]
EW	0.581	0.200	70	-0.36	-16.50	2.16	[0.001]	[0.001]
EW	0.581	1.000	70	-1.62	-16.50	9.82	[0.001]	[0.001]
EW	0.801	1.000	70	-2.06	-16.50	12.48	[0.001]	[0.001]
EW	0.801	1.000	10	-13.70	-16.50	83.05	[0.375]	[0.378]

Table 2. Pseudo-market timing and predictability using other corporate decisions. We simulate 50,000 estimates from the following system of equations:

$$r_t = a + bX_{t-1} + u_t\,,\; X_t = c + dX_{t-1} + v_t$$

where r is an aggregate market return and X is a predictor variable. We use r_t to denote the return in year t and R_{t+k} to denote the k-year cumulative return that starts with the year t return. We consider the following pairs of market returns and predictor variables: stock returns from March of t to February of t+1 and investment plans from Lamont (2000); stock returns and insider trading from Seyhun (1998); the three-year return on long bonds over treasuries and the long-term share of debt issues from Baker, Greenwood, and Wurgler (2003), and the three-year return of payers minus nonpayers and the rate of dividend initiation from Baker and Wurgler (2004). The independent variables are standardized to have zero mean and unit variance. We simulate 100+T values for r and P starting with the actual P_0 and drawing with replacement from the empirical joint distribution of u and v. We throw away the first 100 values, leaving us with a sample size of T. For the three-year returns, we use a sample size of T/3. We use OLS estimates of b for 5,000 separate samples, reporting the average and locating the actual OLS estimate in this simulated distribution with one- and two-tailed p-values. We impose the null hypothesis of no predictability (b = 0) in all cases. The parameters in each panel are as follows. The first row uses the in-sample bias-adjusted OLS estimate for d and the OLS estimate for ρ_{uv}. The second row uses the in-sample OLS estimate for d and sets ρ_{uv} equal to one, leaving the other distributional properties of u and v unchanged. The third row increases the bias-adjusted OLS estimate for d by three OLS standard deviations or to 0.95 (2.7 standard deviations for Panel B and 2.5 standard deviations in Panel C), whichever is lower. The fourth row reduces the sample size to 10.

	Parameter Inputs			Simulation versus Actual Results				
	d	ρ_{uv}	**T**	**Average Coefficient**	**Actual Coefficient**	**Average/ Actual (%)**	**One-Tail p-value**	**Two-Tail p-value**
Panel A. Investment Plans, t = 1948 through 1992, from Lamont (2000)								
r_{mt}	0.088	0.120	45	0.00	-7.15	-0.05	[0.001]	[0.001]
r_{mt}	0.088	1.000	45	-0.36	-7.15	5.08	[0.000]	[0.000]
r_{mt}	0.545	1.000	45	-0.72	-7.15	10.13	[0.001]	[0.001]
r_{mt}	0.545	1.000	10	-3.17	-7.15	44.32	[0.162]	[0.165]
Panel B. Insider Trading, t = 1976 through 1994, from Seyhun (1998)								
r_{mt}	0.358	-0.770	19	1.14	4.99	22.79	[0.105]	[0.116]
r_{mt}	0.358	-1.000	19	1.37	4.99	27.51	[0.111]	[0.115]
r_{mt}	0.950	-1.000	19	2.94	4.99	58.91	[0.182]	[0.182]
r_{mt}	0.950	-1.000	10	5.13	4.99	102.86	[0.469]	[0.470]
Panel C. Long-term Share in Debt Issues, t = 1954 through 1998, from Baker, Greenwood, and Wurgler (1999)								
R_{GLt+3}-R_{GSt+3}	0.616	0.255	15	-0.84	-10.07	8.33	[0.029]	[0.036]
R_{GLt+3}-R_{GSt+3}	0.616	1.000	15	-3.24	-10.07	32.20	[0.063]	[0.063]
R_{GLt+3}-R_{GSt+3}	0.950	1.000	15	-4.92	-10.07	48.89	[0.112]	[0.112]
R_{GLt+3}-R_{GSt+3}	0.950	1.000	10	-6.87	-10.07	68.21	[0.245]	[0.246]
Panel D. Dividend Initiations, t = 1964 through 1998, from Baker and Wurgler (1999)								
R_{Dt+3}-R_{NDt+3}	0.510	0.354	12	-2.35	-24.76	9.48	[0.044]	[0.063]
R_{Dt+3}-R_{NDt+3}	0.510	1.000	12	-7.46	-24.76	30.13	[0.059]	[0.059]
R_{Dt+3}-R_{NDt+3}	0.918	1.000	12	-12.16	-24.76	49.11	[0.112]	[0.112]
R_{Dt+3}-R_{NDt+3}	0.918	1.000	10	-13.81	-24.76	55.79	[0.170]	[0.170]

Table 3. Equity issues and future returns: Before and after January 1975. Regressions of aggregate market returns *r*, either equal- or value-weighted, on one of two measures of past equity issues *E*.

$$r_{mt} = a + bE_{t-1} + u_t$$

The equity share in new issues is the ratio of equity issues to total equity and debt issues. Detrended equity issues is the log difference between the level of equity issues and the average level of equity issues in the previous five years. Both are standardized to have zero mean and unit variance. We report heteroskedasticity-robust t-statistics in braces and the adjusted R^2.

	Equity Share in New Issues = EI / (EI + DI)					Detrended Equity Issues = log (5·EI / (EI_{-1}+EI_{-2}+EI_{-3}+EI_{-4}+EI_{-5}))				
Time Period	**a**	**t(a)**	**b**	**t(b)**	**R^2**	**a**	**t(a)**	**b**	**t(b)**	**R^2**
Panel A. Value-Weighted Returns										
1928-2002	8.10	[3.59]	-6.44	[-3.33]	0.09	9.20	[4.29]	-7.83	[-4.05]	0.15
1928-1974	8.29	[2.60]	-8.18	[-3.59]	0.14	8.16	[2.61]	-7.75	[-3.64]	0.18
1975-2002	9.31	[3.31]	-0.52	[-0.14]	-0.04	11.54	[3.42]	-12.28	[-1.93]	0.05
Panel B. Equal-Weighted Returns										
1928-2002	13.30	[4.01]	-11.92	[-3.49]	0.14	14.75	[4.87]	-16.50	[-3.79]	0.29
1928-1974	14.92	[3.04]	-15.68	[-3.36]	0.21	13.56	[3.01]	-16.42	[-3.48]	0.32
1975-2002	13.66	[3.51]	-0.17	[-0.03]	-0.04	17.42	[3.92]	-21.47	[-1.99]	0.14

Table 4. Out of sample performance: Improvement in Root Mean Squared Error. Root mean squared prediction error from out-of-sample and in-sample predictions. The first three columns report results using a rolling out-of-sample prediction strategy. In the first row, we use all available data up to 1940 and run two regressions of aggregate market returns, either equal- or value-weighted, on a constant and on a constant along with one of two measures of past equity issues. We use the parameters of these two regressions and the 1940 value of equity issues to form two predictions for 1941. We repeat this process for 1942-2002. We report the square root of the average squared prediction error and report the reduction in prediction errors from conditioning on equity issues. In the second row, we start with all available data up to 1950, and so on. The second three columns report results from in-sample prediction. In the first row, we use data from 1941 through 2002 and run two regressions of aggregate market returns, either equal- or value-weighted, on a constant and on a constant along with one of two measures of past equity issues. We use the parameters of these two regressions to form two sets of predictions for 1941-2002. We report the square root of the average squared prediction error and report the reduction in prediction errors from conditioning on equity issues. In the second row, we start with data from 1951 through 2002, and so on.

	Out of Sample			In Sample		
Start Year	**E**	**Constant**	**Difference**	**E**	**Constant**	**Difference**
Panel A. Value-Weighted Returns, Equity Share in New Issues						
1940	17.53	17.93	-0.40	17.19	17.65	-0.46
1950	17.45	17.78	-0.34	17.15	17.57	-0.42
1960	16.87	17.20	-0.32	16.37	16.93	-0.56
1970	17.49	17.67	-0.18	17.04	17.47	-0.42
1980	17.40	16.15	1.24	16.00	16.00	0.00
1990	18.83	18.16	0.67	17.89	17.98	-0.09
Panel B. Equal-Weighted Returns, Equity Share in New Issues						
1940	24.49	25.38	-0.89	23.72	24.93	-1.20
1950	23.99	24.63	-0.64	23.20	24.14	-0.94
1960	24.33	24.79	-0.45	23.18	24.24	-1.06
1970	23.34	22.87	0.48	21.70	22.29	-0.58
1980	23.84	19.52	4.32	18.74	18.76	-0.02
1990	20.48	18.58	1.90	18.30	18.44	-0.14
Panel C. Value-Weighted Returns, Detrended Equity Issues						
1940	17.39	17.93	-0.54	16.69	17.65	-0.96
1950	17.27	17.78	-0.51	16.83	17.57	-0.74
1960	16.48	17.20	-0.72	16.03	16.93	-0.90
1970	17.22	17.67	-0.46	16.98	17.47	-0.48
1980	15.75	16.15	-0.40	15.45	16.00	-0.56
1990	17.64	18.16	-0.51	17.24	17.98	-0.74
Panel D. Equal-Weighted Returns, Detrended Equity Issues						
1940	23.72	25.38	-1.65	22.78	24.93	-2.15
1950	23.05	24.63	-1.57	22.54	24.14	-1.60
1960	22.82	24.79	-1.96	22.26	24.24	-1.99
1970	21.08	22.87	-1.79	20.69	22.29	-1.60
1980	18.51	19.52	-1.01	18.17	18.76	-0.59
1990	16.93	18.58	-1.65	16.39	18.44	-2.05