Unifying Underreaction Anomalies

Andrew Jackson and Timothy Johnson[1]

May, 2002

[1]London Business School. Please address comments to *tjohnson@london.edu.* We are grateful to Stefan Nagel, Allen Poteshman, and Tuomo Voulteenaho for thoughtful comments. We also thank seminar participants at LBS and the University of Illinois. We gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings forecast data, available through the Institutional Brokers Estimate System.

Abstract

This paper asks whether momentum and post-event drift are manifestations of the same underlying mechanism or whether they are separate phenomena. We find that both effects can be attributed to persistence in returns following news which affects expected earnings or earnings growth. Holding these quantities fixed, there is no momentum effect, nor is there post-event drift for our sample of events, which includes seasoned equity offerings, re-purchases, equity-financed mergers, and dividend initiations and omissions.

1 Introduction

Two of the most convincingly documented stock market anomalies are momentum and post-event drift. Momentum refers to the persistent excess returns of winner portfolios over loser portfolios, where winners and losers are judged over a backward-looking horizon of six months to a year.[1] Post-event drift, by contrast, is the tendency of individual stocks' performances following major corporate news events to persist for long periods in the same direction as the return over a short window – usually one to three days – encompassing the news announcement itself.[2] While the details of these two types of behavior (as well as the statistical issues surrounding their measurement) definitely differ, they share a common intuitive interpretation: markets appear to underreact. Periods of good news are followed by periods of unusually high returns relative to natural benchmarks, with the reverse for bad news. This paper addresses the question of whether, indeed, the two types of anomalies are the same phenomenon, just measured in different ways. That is, we would like to know whether there are two puzzles here, or only one.

The answer is not obvious. Momentum itself does not explain post-event drift. For example, firms undertaking share repurchases have usually had negative recent performance (even including the positive announcement return), and yet subsequently outperform. Conversely, corporate news events do not happen all that often, whereas momentum effects seem to be present in stocks generally in all time periods. But perhaps our understanding of what constitutes an "event" is too narrow. Certainly researchers have focussed on those that are most conspicuous and easy to isolate. It could be that drift happens after small and hard-to-observe events as well.

In other words, the hypothesis we wish to test is that momentum is the aggregate effect of post-event drift across all classes of event. If, in fact, future returns are high following any type of good news generally, then presumably, a portfolio of the best performing stocks over the last six months – which ought to contain the firms with the most good news – will the also include the best subsequent performers. Mechanically,

[1]The original momentum findings are in Jegadeesh and Titman (1993).
[2]See Fama (1998) for a critical summary of the literature.

that seems likely enough. But it would not be sufficient. We would also need to be convinced that these event firms were responsible for all the ensuing outperformance. Or, equivalently, we would need to show that, controlling for the occurrence of news events of known impact, there is no residual momentum effect.

The goal of the paper, then, is somewhat prosaic. We are not attempting to explain either of the two anomalies, nor to take a stand on the rationality of the observed patterns. Moreover, while there are still econometric grounds to dispute the patterns themselves, we take both as established fact. Our aim is just to contribute something to the organizational side of the seemingly ever-broadening literature on return anomalies.

Indeed, the most straightforward way, conceptually, to carry out our investigation would be to undertake a giant book-keeping exercise. If we had an exhaustive list of the occurrences of all types of news events, along with estimates of the return drift associated with each, it would be a simple matter to determine if there were an independent momentum effect. Recently, Chan (2001) has done something like this for a sample of 1557 stocks. He identifies all news events for each over the period 1980-1999, and demonstrates both that there is a generic post-event drift – irrespective of news type – and that, among stocks without any news, there is no momentum effect in the sense that, within this group, winners no longer outperform losers in the post-formation period.[3] These results strongly suggest the conjecture of a single underlying effect may be correct. The remaining step is to actually show the connection between the news events (or their absence) and the subsequent return dynamics for each stock.

Broadly, that is our goal here. Our aim is to be able to write down a tractable specification of the the cross-section in terms of some observable indicators of the occurrence of important news. The hope is that, rather than having to describe each firm's expected returns by reference to some omnibus event table, we may be able to find instrumental variables that signal the occurrence of whatever type of news

[3]Pritamani and Singal (2001) also find return continuation (at a 20 day horizon) for news events generically classified as positive or negative.

has persistent effects on returns. If a candidate (or set of them) can both explain momentum (in the sense of knocking past returns out of cross-sectional regressions of current returns) and also account for diverse post-event returns, then there are legitimate grounds for viewing the instruments as identifying a common channel through which both effects operate.

We claim to be reasonably successful in this pursuit. We do, in fact, isolate such explanatory variables. Moreover, in doing so, we move beyond our initial objective of drawing a mechanical connection between the two underreaction puzzles. The variables that explain them both are directly related to changes in expected earnings.

To summarize, our findings are that both momentum and the drift observed after several previously-studied types of event are coincident with changes in either expected earnings or expected earnings growth, and that changes in the latter two quantities account for all the subsequently observed abnormal performance. Holding these expectations fixed, there is no momentum effect, nor is there post-event drift for our sample of corporate actions.

Methodologically, the key step in our analysis is coming up with clean proxies for these changes in expectations. We start from analyst forecast revisions, and then explicitly model and correct for their predictable component. In doing so, we also document a consistent pattern of analyst underreaction both to price changes (which is well-known) and to corporate events (which is less so).

The correspondence between patterns of analyst underreaction and price underreaction is highly suggestive of a behavioral explanation for return persistence. However, rational stories based on contemporaneous shocks to cash-flow risk may also be consistent with these patterns. In short, we do not yet know why news that affects earnings expectations has persistent effects on returns. But we do think that identifying this association considerably sharpens the focus of the theoretical effort to understand what is going on.

Our work is closely related to two other important recent papers. Chan, Jegadeesh, and Lakonishok (1996) sought to clarify the relationship between momentum and post-earnings-announcement drift. Their goals were similar to ours, and, indeed, some of our techniques and variable definitions are directly attributable to them. Ultimately they conclude that momentum is not fully explainable by changes in earnings expectations or earnings surprises (although these do account for a significant fraction of the effect). Nor does momentum explain the drift following these surprises. Building on their results, we employ what we think are improved measures of expectations, and claim that these can finish the job.[4]

In a similar vein, Vuolteenaho (2002) runs firm-level panel vector autoregressions of changes in future cash flows and returns. He finds that after controlling for shocks to future cash flows there remains a residual momentum effect, especially for smaller firms. This is due to positive correlation between cash flow shocks and expected return shocks that can be attributed either to underreaction or risk. Our innovation with respect to his set-up is to control for changes in expected future earnings directly via observing revisions to analysts forecasts rather than indirectly through an accounting-based present value formula. While we must sacrifice the elegance of the VAR framework, our specification may be viewed as an extension of his.

The outline of the paper is as follows. The next section describes our model and relates it to similar formulations in some other, well-known settings in financial econometrics. Section 3 fits the basic model and demonstrates that it explains momentum. In Section 4 we examine a set of corporate events that have been widely used in the long-horizon event-study literature. An expanded version of our original model is then shown to explain the abnormal returns subsequent to those. Section 5 contains some concluding remarks on the interpretation of our findings.

2 The Model

To investigate the hypothesis that underreaction anomalies can be traced to changes in fundamental expectations, one would like to simply run regressions of returns on

[4]We plan to examine post earnings announcements drift in subsequent work, but do not do so here.

proxies for those expectations alongside other explanatory variables known to be associated with return persistence. Things become less straightforward, however, when the available expectational proxies are poor. In this study, we use changes in consensus earnings forecasts (and growth forecasts) from IBES, which turn out to be both autocorrelated and forecastable from other predictors, including momentum. In this section we derive the return specification that ensues from cleaning up these "expectations."

Before launching into the equations, let us describe what needs to be done. The defining property of true conditional expectations is non-forecastability of their changes. So the first step in cleaning up our proxies is to remove their predictable component. But there is a second step as well. Even if there is no innovation in the proxy at time t, if there are contemporaneous changes in other variables that imply future changes in the proxy, then the true conditional expectations at t must be updated to reflect this. To make this concrete, if we know that the analysts polled by IBES do not weigh current returns sufficiently in changing their growth estimates, then even when these estimates are unchanged today, if observed stock returns cue us to the updates they will subsequently have to make, then we will impound those updates in our true expectations at the time of the return observation. We need our cleaned-up expectations to correctly reflect future predictable revisions immediately.

Mathematically, suppose $i^\star$ is a vector of parameters that affects a firm's value, and that $\{i_t\}$ is any (noisy) series that eventually converges to $i^\star$. Then, $\mathrm{E_t}[i^\star] = i_t + \mathrm{E_t}[\sum_{k=0}^{\infty} \Delta i_{t+k}] = i_t + \sum_{k=0}^{\infty} \mathrm{E_t}[\Delta i_{t+k}]$, (assuming the latter is finite). The correct measure of the amount of news about $i^\star$ that arrives at time t then is

$$\xi_t \equiv \mathrm{E_t}[i^\star] - \mathrm{E_{t-1}}[i^\star] = \sum_{k=0}^{\infty} \left[\mathrm{E_t}[\Delta i_{t+k}] - \mathrm{E_{t-1}}[\Delta i_{t+k}]\right]. \tag{1}$$

Equation (1) is quite general: it does not actually require the true parameter to be constant through time. Moreover, it gives the correct expression for the change in conditional expectations even if the estimates never do, in fact, converge to the true value, as long as any bias component is stationary. In our setting, we interpret the observed IBES estimates as the noisy series $\{i_t\}$, with Δi_t being the time-t revisions of those numbers.

To construct the true forecast innovation series, then, we require a model of the evolution of Δi_t. We return to this in a moment. But first let us write down our basic hypothesis using the current notation. If r_t is the excess return at time t (for a given asset), then our model simply says

$$r_t = \alpha_0 + \alpha_1' \, \xi_{t-1} + \epsilon_t^{(0)}. \tag{2}$$

That is, innovations in our fundamental variables cause changes in expected returns, and other things do not.[5] The basic model (2) envisions only one-period-ahead effects (so $\mathrm{E_t}[r_{t+2}]$ is a constant), which of course is not essential. We have also fit two-lag and geometric lag models. And the best specification will, in general, depend on the time interval Δt. While we intend to report refinements based on these variations in future work, we do not want to obscure the fact that the economic content of what we are doing boils down to this one equation.

Econometrically, (2) complicates matters because it turns out that returns also predict changes in Δi.[6] At this stage, let us suppose for simplicity, that they are the only quantity – besides lags of Δi – that does so. This implies a specification like

$$\Delta i_t = \mu + A(L)\Delta i_{t-1} + a(L) r_{t-1} + \epsilon_t^{(1)} \tag{3}$$

where μ is a constant and A and a are polynomials (of appropriate dimension) in the lag operator L. Equation (3) is just a linear specificatino in terms of observable quantities. And, given its parameters, we can compute the pure forecast innovation series ξ. However the computation will also depend on the parameters in (2) because expectations of future returns are involved.

Despite this complication, the resulting expression for the true change in condi-

[5]We might add the hypothesis that $\alpha_1 > 0$, when positive changes in ξ denote good news, to describe the effect as persistence and not reversal.

[6]To our knowledge, this was first documented by Abarbanell (1991).

tional expectations can be written compactly as

$$\xi_t = (I + EL)^{-1} F \lambda_t \tag{4}$$

where

$$F = [I - A(1) - a(1)\alpha_1']^{-1} \tag{5}$$

$$E = F\, a(1)\, \alpha_1' \tag{6}$$

$$\lambda_t = (\epsilon_t^{(1)} + a(1) r_t). \tag{7}$$

(The derivation appears in the appendix.) Combining (4) with (2), our return model becomes

$$r_t = \alpha_0 + \alpha_1'(I + EL)^{-1} F \lambda_{t-1} + \epsilon_t^{(0)}. \tag{8}$$

Intuitively, equation (8) is just a regression of returns on a filtered version of the series $\{\lambda_t\}$. That series is the augmentation described above of the current unexpected change in the observed forecast by the cumulative future revision predicted by current returns. There is a further filtering step required (implemented by the operator $(I + EL)^{-1}$) only because of the added complication that the return part of λ_t is itself partially predictable according to (2). With typical parameter values, eigenvalues of E are small, and the contribution of the lagged terms is of secondary importance.

While (8) may be understood as just a linear regression on the corrected innovations, estimation is rendered harder by the fact that the regression coefficient α_1 (which gives the return premium associated with the innovation ξ) is also involved in the matrices E and F which define ξ. So, even though we have a straightforward linear time-series model for Δi, this ends up producing a non-linear model for returns. This means that we will not ultimately be able to gauge our model by the familiar method of running OLS regressions of returns on some new candidate variables alongside established competitors. Instead, we first estimate (3), then fit the model (8) (via non-linear least squares), and then evaluate it by regressing its residuals on the competing predictors.

Since this two-step procedure involves using some of the same predictors (notably, past returns) in the Δi equation that we are claiming to absorb in our specification of r, we thought it would be useful, in interpreting our results, to draw some parallels with similar, familiar procedures from other econometric settings.

First, our approach is closely analagous to testing the permanent income hypothesis by initially specifying a time-series model of income, and using its estimates to identify the long-run effect (or permanent component) of observed innovations, which is not directly observable.[7] This variable, call it X, is analagous to our ξ. It is then put on the right-hand side of a consumption regression, along with income and other covariates. We regard the permanent income hypothesis as supported if the current income term no longer enters, though it is, of course, present implicitly in the synthetic variable X. The statement then is not that income shocks do not affect current consumption, but that they do so only to the extent that would be expected based on separate, independent estimates of what their likely impact on the true determinant, X, should be. This is precisely what we will conclude about the role of past returns in predicting future returns.

A second useful comparison comes from the recent work in empirical corporate finance which tests for the sensitivity of firms' investment to current cash flow. Here the theoretical hypothesis is that current investment opportunities alone should explain current investment. Since opportunities are unobservable, one strand of the literature takes a structural approach, estimating profitability via instrumental variables, using the fitted model to solve for the expected present value of future profits, and from this deducing the true "marginal Q", or return on investment.[8] Again, this construction is much like our computation of ξ. The subject under study is whether the synthetic Q alone can explain firms' current investment, or whether current cash flow also matters. But current cash flow (like past returns in our set-up) is also one of the most important instruments used in the first-stage structural model. Still, we

[7]See Quah (1990) for the development of this idea. It has recently been implemented by Falk and Lee (1998).

[8]This methodology is used by Gilchrist and Himmelberg (1995) following Abel and Blanchard (1986).

regard the null as supported if cash flow only enters the investment model via this constrained channel (it doesn't).

With these parallels in mind, our methodology can be viewed as a standard structural identification scheme in estimating simultaneous equations with errors in variables. The important points to bear in mind are that we would be misspecifying the Δi dynamics if we did not include past returns (and other predictors) which play a highly significant role statistically; and also that, since Δi is fit independently of the return model, the predictors are not subsequently entering the return model as free parameters. Instead, their coefficients are highly constrained, and we effectively tighten them further by neglecting the estimation error from the first stage, treating the point estimates as known constants in the second stage. Finally, we will explicitly test the model's restrictions by, in effect, giving the data a chance to un-do the constraints.

The next section introduces the data series we employ, and describes how we carry out the estimation and testing steps. Results are presented for the basic model described by equations (2) and (3). From these we are able to conclude that our corrected informational variables do, in fact, account for momentum effects. After that, we turn to the connection to post-event drift in Section 4.

3 Initial Results

As discussed at the start of the preceding section, the initial expectational variables employed in this study are based on IBES consensus earnings forecasts. Our working hypothesis is that good news generates persistent positive abnormal returns if and only if the IBES forecasts – corrected as described above – concurrently increase, with negative abnormal returns likewise when they decrease. For each stock tracked by IBES, the monthly data records include average forecasts (by varying numbers of analysts) for a number of fiscal horizons, as well as a separate consensus forecast for five-year growth. One could imagine different types of news affecting earnings at different horizons independently, and hence an informational role for changes in each horizon's forecast. For parsimony, we employ only revisions to the nearest fiscal year's estimate and to the five-year growth number. The latter is included to capture news about long-term cash-flow prospects, and is at least partially motivated by models such as that in Johnson (2002), which suggests an independent role for growth-rate risk.

The basic time unit of the study will be six month intervals. Thus our fundamental variable Δi_t consists of two components for each stock (which we call DFY1 and DLTG in the tables), defined to be (respectively) the change over the preceding six months in the mean forecast for the next fiscal year[9], and the change in the mean long-term growth forecast over the same period. Although the long-term forecasts do not go back in time as far as the annual ones, we still have observations for both series for several thousand stocks for each month from the beginning of 1983 through the end of 1999. Following Chan, Jegadeesh, and Lakonishok (1996) and others in this literature, we further transform each variable into its percentile rank in each month. Table 1 summarizes some aggregate properties of the sample.

The first step in our estimation is to fit the prediction equation (3) for our bivariate process Δi in order to extract its innovation component. Table 2 shows estimates of a specification which incorporates one lag of Δi itself and two lags of returns.[10] The estimates shown are time-series averages of the coefficients from monthly cross-sectional regressions (a la Fama and MacBeth (1973)), with standard errors corrected for the induced six-lag autocorrelation.

The most notable finding here is the extreme predictability of the changes in annual forecasts. It is well known in the accounting literature that analysts tend to smooth their revisions over time, and thus the significance of lagged revisions is not surprising. The finding that near-term forecasts underreact to stock returns at lags up to one year is consistent with the results of Abarbanell (1991). The same is true of revisions to long-run growth forecasts. However, accounting for the momentum

[9]A complete description of all data fields and variables is given in Appendix B.
[10]Including longer lags has no significant effects on the results below.

Table 1: *Stocks With IBES Data*

Year:	Number of Firms	Average size decile: mean	median	Stocks by Exchange: NYSE	AMEX	NASDAQ
1983	1546	4.1	3.0	1140	107	299
1984	1715	3.4	2.0	1140	119	456
1985	1945	3.3	1.0	1141	137	666
1986	1963	3.3	2.0	1139	135	688
1987	2007	3.2	2.0	1113	144	749
1988	1982	2.7	1.0	1005	139	838
1989	1979	2.8	1.0	1026	142	811
1990	2058	2.8	1.0	1020	157	881
1991	2027	2.9	2.0	1009	133	885
1992	2050	3.2	2.0	1070	120	859
1993	2181	3.4	2.0	1136	114	930
1994	2452	3.6	2.0	1227	97	1126
1995	2638	3.8	3.0	1278	82	1275
1996	2791	4.0	3.0	1362	74	1353
1997	3069	4.0	3.0	1456	80	1531
1998	3355	4.2	3.0	1568	89	1697
1999	3435	4.6	4.0	1621	101	1712

The table describes the stocks for which there are IBES forecasts for both next fiscal year's earnings and long-term (five year) growth. For a firm to be included in a given month it must have forecasts for at least the previous six months, and have stock prices on CRSP for the previous twelve months and the following six months. Size deciles are computed using month-end NYSE breakpoints.

dependency, analysts actually over adjust this number in the sense that there is a significant tendency to reverse the direction of time-t changes at $t+1$. This overreaction has been previously documented by LaPorta (1996).

The next stage in our program involves estimating the parameters of the return equation. This requires as inputs the coefficients of the Δi evolution equation, for which we use the point estimates in Table 2.[11] Given these values for A and a, and the fitted Δi residuals $\hat{\epsilon}^{(2)}$, the only unknowns in equation (8) are the αs. We fit

[11] By not accounting for the estimation uncertainty in these, we overstate the accuracy of our estimate of the return parameter α_1 below. On the other hand, when we test our return residuals for remaining underreaction effects, the t-statistics overstate significance levels, i.e. we bias the tests in the direction of rejecting our model.

Table 2: *Forecasting Forecast Revisions*

	μ	DFY1_{t-1}	DLTG_{t-1}	R6	LR6
DFY1_t	.0021	.2718	-.0012	.3219	.0924
	(1.60)	(38.03)	(0.43)	(20.51)	(9.95)
DLTG_t	.0006	-.0071	-.0573	.2052	.0744
	(0.60)	(1.21)	(11.07)	(23.02)	(9.45)

The table shows estimates for the semi-annual specification: $\Delta i_t = \mu + A\Delta i_{t-1} + a_1 r_{t-1} + a_2 r_{t-2} + \epsilon_t^{(2)}$ where Δi is the bi-variate process of IBES forecast revisions whose components are DFY1 and DLTG, the percentile rank changes in next-fical-year and long-term growth respectively. R6 and LR6 are the stock returns over months $t-5$ to t and $t-11$ to $t-6$ respectively. Coefficients are estimated by Fama-MacBeth regressions for each component. The data are monthly observations from 1983 through 1999. Standard errors are adjusted for induced six-month autocorrelation. The resulting t statisitics are in parentheses.

these by non-linear least squares with a time-varying intercept. The resulting pair of return premia, the two components of α_1, is shown in Table 3.[12]

Table 3: *News Premia Estimates*

The first column of the table gives return premium estimates, α_1, in the return model $r_t = \alpha_0 + \alpha_1' \xi_{t-1} + \epsilon_t^{(1)}$, where ξ is the corrected news process defined by equations (4) - (8). The parameters are estimated by non-linear least squares (see footnote 12). Estimates of the ξ process, which depend on α_1, are then computed for each stock in our sample. The two right-hand columns report the average cross-sectional (CS) and time-series (TS) standard deviations of these imputed series.

			avg std dev:	
	α_1	(t stat)	CS	TS
ξ_1	.0514	(3.03)	.351	.377
ξ_2	.0159	(2.15)	.237	.246

Innovations to near-term expected earnings have about three times the effect that innovations to long-run growth do upon expected returns. For reference, the table also shows the average cross-sectional and time-series variability[13] of the two components

[12]Standard errors are computed from the asymptotic covariance matrix $T^{-1} A^{-1} B A^{-1}$ where B is the estimated covariance of the normalized scores of the NLLS criterion function, Q, and A is Q's numerical second derivative. See Amemiya (1985) Theorem 4.1.3.

[13]These are arithmetic averages of standard deviations. Root mean squares were similar. The

of the innovation series. Combined with the α estimates, these indicate that the magnitude of the fitted model's effects are of the order of plus and minus 250 basis points of return differential (in expected six-month returns), with, again, the earnings revisions accounting for a larger share than the growth rate revisions.

Table 4 shows results from ordinary Fama-MacBeth regressions using momentum (trailing six-month returns) as a predictor. The first panel shows results when the raw returns themselves (over the next six months) are the dependent variable. The second panel uses instead the residual part of these returns in excess of that fitted by our model. The message is clear. Regardless of what other controls are in the regression, the momentum anomaly is strongly present in our sample of returns, but has been almost entirely erased by our model.

It is important to point out that, although past returns are indeed incorporated in our innovation variable λ, our model was not free to adjust their weight. In this regard, the specification in Table 4 which also includes the unadjusted IBES variables is especially important. If all our model was doing was picking up the momentum component of λ, the story would come out here. The data would then take the opportunity to undo the spurious inclusion of the IBES variables by negatively weighting them, while simultaneously according momentum its "true" weight. This is not what happens. Instead, all the predictors in the residual regression are insignificant. This suggests that the model has not simply included momentum by the backdoor.

To make the point more formally, we present results of two specification tests in table 5. In the previous section, we submitted that our model might best be understood as a restricted linear specification of returns (in terms of their own lags and other predictors), where the restrictions are imposed from the analyst revision equation. With that in mind, we test whether loosening these restrictions significantly improves the fit of the model via a Lagrange multiplier-type statistic. Specifically, this test envisions the alternative hypothesis as allowing the coefficients on R6, DFY1, and DLTG to be unrestricted, but constant across firms.[14] Under the null, this statistic

average time-series correlation between the two components is 0.35.

[14]The statistic is computed as $g_T' \hat{W} g_T$ where $g_T = \frac{1}{T}\sum h_t, h_t = \hat{u}_t' X_t, \hat{u}$ are the residuals, X_t is

Table 4: *Return Regressions*

R6	DFY1	DLTG	E/P	SIZE
DEPENDENT VARIABLE=6-MONTH RETURNS				
.0577				
(5.36)				
.0600	.0606	.0135		
(3.96)	(5.26)	(3.75)		
.0624	.0529	.0147	.0238	.0096
(5.03)	(5.87)	(4.21)	(1.03)	(0.61)
DEPENDENT VARIABLE=MODEL RESIDUALS				
.0178				
(1.07)				
.0127	.0088	.0007		
(0.80)	(0.75)	(0.18)		
.0155	-.0029	.0019	.0264	.0148
(1.20)	(0.14)	(0.51)	(1.16)	(0.91)

The table shows results from Fama-MacBeth regressions for monthly stock returns (top panel) and estimated model return residuals (bottom panel), for different specifications of independent variables. R6, DFY1, and DLTG are the six-month trailing returns, and IBES forecast series defined in the text. E/P is the percentile rank of earnings to price, where the numerator is the consensus forecast for the next fiscal year. SIZE is the percentile rank of market capitalization. The data are monthly observations from 1983 through 1999. Standard errors are adjusted for induced six-month autocorrelation. The resulting t statisitics are in parentheses.

is asympotically distributed as $\chi^2(3)$, and is also equivalent asymptotically to the sample size times the R^2 from a pooled regression of the model residuals on these instruments, corrected for overlapping observations. A similar test may be based on the coefficients in the cross-sectional regressions of the residuals on the instruments. Here, we can construct a another $\chi^2(3)$ statistic to test the restriction that these are jointly zero. This is analogouos to the usual F statistic, corrected for the induced autocorrelation of the cross-sectional coefficients. We label this the Fama-MacBeth statistic in the table.

Neither test is able to reject the null hypothesis that our constraints hold. To be careful, our conclusion is not that there are no momentum effects in returns, but that past returns affect future returns only to the degree that would be expected, given their role in proxying for future revisions to earnings forecasts.

Table 5: *Tests of Model Constraints*

The table shows the results of two specification tests based on the fitted model residuals. The Lagrange multiplier statistic weights the cross-sectional average product of the residuals and the instruments (See footnote 14.) The Fama-MacBeth statistic tests the hypothesis that the cross-sectional regression coefficients of the residulas on the instruments are jointly equal to zero. Both statistics are asymptotically distributed as $\chi^2(3)$. Asymptotic p values are in parentheses.

Instruments:	Lagrange Multiplier	Fama-MacBeth
R6, DFY1, DLTG	1.5058 (0.68)	1.6336 (0.65)

In the next section we extend these results to encompass the return persistence that follows corporate news events. Our basic framework for estimation and testing will be the same as that used in this section. But we will have to extend the econometric specification of Section 2 slightly.

the matrix of derivatives of the mean specification with respect to the constrained variables, and $\hat{W}$ in an estimator of the asymptotic covariance matrix of the h process. This may also be viewed as a GMM test of orthogonality in cross-sectional mean of instruments and residuals.

4 Post-Event Drift

Our inferential approach to post-event return persistence is based on simple calendar-time regressions, and thus departs somewhat from the usual methodology of the long-horizon event study literature. Specifically, the procedure is as follows. For a given type of corporate event, we include a dummy variable in our Fama-MacBeth regressions which is one for firm i at time t if the announcement date of such an event occurred for firm i in the six months prior to (and including) t. If the dependent variable in the regression is the return in the following six months, then we are implicitly examining the average post-event drifts for months +1 to +12 in event-time. This set-up allows us to employ the same methodology used in the last section, as well as to include whatever controls may be of interest in determining the appropriate benchmark.

To find out if our corrected earnings expectation variables can explain post-event drift, we collect samples for three types of event that have been examined extensively: secondary equity offerings (SEOs), share repurchases, and stock-financed mergers.[15] We collect announcement dates on all three from SDC for the years for which we have constructed our expectation variables.[16] Table 6 summarizes the events by year for those that also intersect with our IBES sample.

The table confirms that the pattern of underreaction documented by previous studies is present in our data. For SEOs and share repurchases the sign of the average excess three-day announcement returns is the same as that of returns for the following 12 months (starting with the first post-announcement month). For stock financed mergers, the initial news is ambiguous. The mean announcement return is small and positive and the median is negative, which probably indicates that we haven't identified the real event-date correctly in a lot of cases.[17] In any case, the one year

[15]Some of the primary work on these can be found in Asquith and Mullins (1986), Loughran and Ritter (1995) and Speiss and Affleck-Graves (1995) (for SEOs), Ikenberry, Lakonishok, and Vermaelen (1995) (for buybacks), and Travlos (1987) and Loughran and Vijh (1997) (for mergers). Collectively, these were also the basis for the study of long-horizon inference in Mitchell and Stafford (2000), and the study of underreaction by Kadiyala and Rau (2001).

[16]See Appendix B for details on our selection criteria and definition of events.

[17]We don't use the announcement returns for anything below, so this is immaterial.

returns are reliably negative. When we perform the return regressions described above, the t-statistics for the event dummies are all significant in most specifications.[18]

The bad news for the model of the last section is that it does not eliminate all of this underreaction. In regressions of its residuals on the event dummies (which we omit) the coefficient on stock-financed mergers does become insignificant. But the SEO coefficient is only lowered slightly in magnitude, and the share-repurchase one actually increases. This apparently indicates that these events lead to return drift beyond that attributable to the contemporaneous changes they cause in earnings expectations.

The picture changes fundamentally, however, when, in addition to the contemporaneous correlation, lagged correlations are taken into account. It turns out that analyst forecasts also drift in periods following our events. Or, equivalently, lagged values of the event dummies enter significantly in Δi regressions. Statistically, this means equation (3) is now mis-specified (and the derivation of the subsequent model equations needs to be modified). Intuitively, it means that the corrected expectational innovations (our ξ) ought to react to the occurrence of one of our events so as to incorporate all future predictable revisions in the uncorrected series. That is, we must un-do the underreaction of the analysts.

To see how this works mathematically, write the amended version of (3) as

$$\Delta i_t = \mu + A(L)\Delta i_{t-1} + a(L)r_{t-1} + c(L)\Delta x_{t-1} + \epsilon_t^{(1)} \tag{9}$$

where now Δx_t is the event indicator process. (For the exposition we will assume there is only one type of event. In the version we fit there will be a separate process, and a separate lag polynomial, for each of our three types.) Next, assume x_t is a martingale. This is also for simplicity, but it is not far off.[19] Then, as shown

[18]We have also used a design where the dummy variable in the regressions is set equal to the sign of the (three-day) announcement return. The only important difference in our results with this alteration is that the drift following all-stock mergers fails to be significant with either the returns or the model residuals.

[19]In fact, previous buy-backs tend to predict buy-backs, and momentum predicts SEOs. But the effects are weak. Correcting for this slightly strengthens our results.

in Appendix A, equations (4), (5), and (6) are formally unchanged. But now (7) becomes

$$\lambda_t = \epsilon_t^{(1)} + a(1)\, r_t + c(1)\,(\Delta x_t - \mathrm{E}[\Delta x_t]). \tag{10}$$

As with the return term, the event correction now involves the sum of all the regression coefficients in (9) multiplied by the current value of the indicator. Table 7 shows our estimates of these coefficient sums. For equity-financed mergers, we do not detect any significant analyst underreaction. But for buy-backs and SEOs we find significant coefficients, nearly all of the same sign, out to lags of nearly six years for revisions to both year-one estimates and long-term growth forecasts. This finding, though ancillary to the main purpose of our study, is nevertheless remarkable in its own right. We have documented a long-horizon post-event drift in analysts' forecasts which parallels closely the pattern in returns.[20]

Continuing now with our program, incorporating the lags of the event indicators in our Δi specification, we use the new estimates of the parameters in (9) to again construct the innovation series λ_t.[21] As before, we then use λ_t to simultaneously estimate the return response coefficients α and build the corrected news process ξ_t. Residuals from our return model are then tested for remaining predictability. The results are shown in Table 8.

The first panel shows various specifications for the raw returns themselves, and demonstrates the significance of the post-event drift. Interestingly, the post-merger drift is almost all accounted for by the raw IBES revisions, whereas the other two event types are not. The fact that analysts revisions also seem to respond correctly to mergers, but not to buybacks and SEOs, strongly suggests that our model is indeed isolating the true informational component of the forecast series.

The second panel of the table runs the same regressions with the model residuals. All the event dummies – as well as momentum itself again – are now reduced to

[20]The predictable downgrading of firms after SEOs is studied in Teoh and Wong (2002) and Dechow, Hutton, and Sloan (1999). We are not aware of similar work on repurchasing firms.

[21]Because of the long lags required of the event variables, this series now starts in 1985. Neither the new estimates of A, a, and α, nor their standard errors differ appreciably from the values reported in Section 3.

levels indistinguishable from zero. Once again, we qualify our conclusion due to the presence of some of the event variables implicitly on the right-hand side. We do not find that there is no post-event drift, but rather that there is as much as would be expected if the only role of the event variables was to instrument for future forecast revisions.

As a further check on our ability to explain previously reported results, we next turn to the case of firms initiating or omitting dividends. Michaely, Thaler, and Womack (1995) report significantly negative announcement returns for omitters followed by excess negative returns – even after correcting for momentum – for the next several years. Initiators had the opposite pattern, though somewhat weaker. We gather initiation and omission information from CRSP for our sample stocks (as described in Appendix B), and test whether indicators of these events are significant in return regressions after subtracting our model's expected component, as above.

Stochastically, there is an important difference between the event indicators for dividend changes and those for the corporate actions used above. The former cannot be independent through time. This is just definitional: a firm that has previously initiated dividends cannot do so again (without first omitting), likewise an omission can only follow an earlier initiation, not another omission. If S_t is an indicator equal to one if a firm is a current divident payer at time t, and zero if it is not, then S is a first-order Markov process. Its transitions, ΔS, are +1 for initiations and -1 for omisisons, which we model as

$$\Delta S_t = \mu^{(2)} + A^{(2)}(L)\,\Delta i_{t-1} + B^{(2)}(L)\,S_{t-1} + a^{(2)}(L)\,r_{t-1} + \epsilon_t^{(2)}$$

to capture the level dependency, as well as the tendency for past earnings growth and returns to predict dividend changes.

This new specification requires another solution of our model, which is again worked out in Appendix A. There is a surprising feature of the solution this time: changes in dividends drop out entirely. This is entirely an artifact of modeling S as a stationary process on an infinite horizon. In essence, in evaluating the expected future earnings revisions following a dividend change, the model takes account of not just the current event but all future dividend events as well. Since, under the model, every initiation is eventually followed by an omission, the total sum of all future revisions is always zero. While we could improve on this bit of artificiality by explicitly modeling the finite lifetime of firms, we choose not to do so because, if anything, sticking with our formalism handicaps us in this case. It does turn out that there is some underreaction in earnings revisions to dividend events (that is ΔS enters the Δi equation). But now this is not incorporated in the definition of ξ, and so does not help explain any underreaction in returns.

We do have to rebuild our residual series, however, because we get slightly different point estimates on the other model variables when ΔS is incorporated as a predictor. Then, as before, we regress raw returns and residuals on all the underreaction variables (events, and momentum) including dividend changes. The results are shown in Table 9.

The first panel confirms the results of Michaely, Thaler, and Womack (1995). This is important because our sample is mostly disjoint from theirs (which end in 1988). In our data too there is indeed strong post-event drift, controlling for momentum, and omissions are more significant than initiations.

The second panel shows that these events too do not imply unexpected excess returns when controlling for changes in earnings expectations. Whether initiations and omissions are considered separately or together, whether the other corporate events are included or not, dividend changes do not explain any of the residual variation in returns not already captured by our model.[22]

Table 10 presents the results of the specification tests introduced in the previous section. Recall that these tests assess the degree to which the model would be improved by relaxing the constraints imposed on the components of our constructed ξ variable. Since none of the tests comes close to rejecting, there seems little reason to fear that the forecast revision model has somehow merely re-arranged the predictor

[22]The same holds true with other controls in the regression. As in the previous tables, these specifications are not markedly different, and so we omit them for brevity.

variables in an arbitrary fashion. We also note that these tests are now more powerful than the ones in Section 3, due to the added instruments. Intuitively, it becomes increasingly implausible that our return model, with only two free parameters, could simulaneously be compensating for more and more false restrictions. And, of course, these are not arbitrary or weak instruments: they are jointly and collectively significantly related to future returns. So the tests are imposing highly stringent conditions on the return specification.

To check that our results are not being driven by small firms or firms poorly covered by analysts, we repeat our procedure on the set of firm-months in which estimates are available from five or more forecasters. A common finding in the anomalies literature is that return patterns are often due largely to small, illiquid stocks. Certainly too, consensus forecast dynamics are likely to vary with the level of coverage. Moreover, Hong, Lim, and Stein (2000) have explicitly connected the level of coverage with the strength of momentum effects. So, in fitting the same coefficients across firms in both our return equation and our forecast revision equation, we are undoubtedly inducing some misspecification. Though it is not obvious why this would make it easier for us to explain the underreaction anomalies, splitting the sample by coverage seems like a sensible robustness check.

Table 11 shows our before and after return regressions for both the high coverage and the low coverage sub-samples. As expected, the high coverage returns (Panel A) are overall less "anomalous", exhibiting smaller and less significant coefficients on all the underreaction variables.[23] Nevertheless, substantial predictability does remain in these firms. The model is no less successful with them, however. Residual returns have again been purged of all significant underreaction effects, exactly as in the full sample, and as in the low coverage group (Panel B). On this basis, we conclude that any misspecification resulting from our pooling of firms cannot be affecting our conclusions.

[23] Interestingly, it turns out the forecast revisions for the better covered firms are *more* predictable. Forecast dynamics are roughly the same across the two groups however.

To summarize, we have presented evidence that strongly supports the view that our cleaned expectational variables, together with the simple return model (2), explain both classes of underreaction phenomenon. It is thus reasonable to infer that both anomalies are manifestations of the same basic mechanism, and do not present separate puzzles. News that causes fundamental earnings expectations to change does have persistent effects on returns. Holding these expectations fixed, neither momentum nor the corporate actions examined here appears to alter expected returns.

5 Concluding Remarks

This paper has argued that return persistence – following either important corporate news or exceptional past returns – occurs if and only if those events coincide with changes in earnings (or earnings growth) expectations. We have not attempted to explain why even this type of fundamental news should produce persistence. But we think the fact that it does advances our understanding and narrows the focus of the search for a theoretical explanation.

We see at least two interpretations of our findings. On the one hand, behavioralists will not fail to note that in building our "corrected" expectational proxies from reported forecasts, we uncover precisely the same biases in the IBES data that are observed in returns themselves: analysts appear to significantly underreact to 6 to 12 months of stock returns and to up to 5 or 6 years of corporate actions. Certainly a natural interpretation of this coincidence is that the analysts' errors were, in fact, the market's errors, and that our adjusted expectational measures, while summarizing the informational content of those errors, are merely a mechanical construct.

On the other hand, if one views our derived variables as legitimately correcting noisy proxies in the direction of actual expectations, the outlook for a purely rational theory of return persistence seems promising. Under this interpretation, if better observations of agents' true expectations were available (or perhaps if analysts forecasts were not distorted by institutional factors), the econometric complexities of our model would unwind (Δi would be the same as ξ in our notation), and we would be left with a much-consolidated ordinary factor representation of the cross-section of expected returns.

Either way, linking the underreaction anomalies to expectations (even if perhaps mistaken ones) about fundamental quantities such as earnings and earnings growth appears to us to cast these puzzles in a much simpler light than they have appeared heretofore. We may perhaps now be able to dispense with the more abstruse and exotic theories that have sometimes seemed necessary to explain them.

Finally, on a basic level, we think our work also bolsters the case for accepting the existence of both of the underreaction anomalies themselves. It is always important to apply rigorous and stringent critiques to reports of new asset pricing "puzzles". But the direct channel we have demonstrated between fundamental variables and both momentum and post-event drift would seem to make it increasingly difficult to maintain the position that either or both are merely a product of data-mining or inferential errors. Both coincide with expectational revisions, and the magnitude of those revisions exactly matches the magnitude needed to account for the subsequent return persistence.

Table 6: *Corporate Events in IBES Sample*

Year:	Number of events		
	SEO	BBK	EFM
1983	1156	156	460
1984	391	435	487
1985	594	210	91
1986	912	238	122
1987	593	830	132
1988	281	339	90
1989	362	600	152
1990	310	903	109
1991	748	402	169
1992	875	534	256
1993	1202	550	329
1994	836	950	411
1995	1007	999	479
1996	1335	1335	544
1997	1118	1195	662
1998	758	1739	631
1999	850	1363	533
Mean event return	-.0069	.0186	.0030
Mean one year return	-.0359	.0215	-.0357

The table describes the corporate action sample used in Section 4. Announcment dates for secondary equity offerings (SEO), share buy-backs (BBK), and equity-financed mergers (EFM) are taken from SDC for all firms for which we have the IBES fields used in Section 3. Also shown are the mean 3-day announcement return across the whole sample, and the mean 12-month return commencing in the first post-announcement month. Both returns numbers are excesses over the S&P500 return for the same period.

Table 7: *Corporate Events as Predictors of Forecast Revisions*

	Event Type		
	SEO	BBK	EFM
Coefficient sum:			
Dependent=DFY1	-0.0958	0.0699	0.0000
Dependent=DLTG	-0.2157	0.0574	0.0000
Lags selected:	11	9	0
χ^2 drop lag :	.0440	.0085	—-
χ^2 add lag :	.0785	.3921	.4470

The table shows the sum of the estimated coefficients on lags of corporate event indicator variables (c.f. equation (9)). Lags are included until χ^2 tests reject the significance of an additional lag at the 95% level. The three event types are secondary equity offerings (SEO), share buy-backs (BBK), and equity-financed mergers (EFM).

Table 8: *Return Regressions*

R6	SEO	BBK	EFM	DFY1	DLTG	E/P	SIZE
Dependent Variable=6-Month Returns							
.0547	-.0174	.0083	-.0201				
(5.10)	(3.05)	(1.40)	(3.47)				
.0595	-.0170	.0110	-.0084	.0644	.0138		
(3.75)	(2.96)	(2.43)	(1.30)	(5.23)	(3.51)		
.0602	-.0169	.0085	-.0098	.0611	.0137	.0056	.0095
(4.65)	(3.18)	(2.62)	(1.53)	(7.03)	(3.71)	(0.23)	(0.55)
Dependent Variable=Model Residuals							
.0200	-.0016	.0053	-.0038				
(1.14)	(0.26)	(1.18)	(0.54)				
.0140	-.0033	.0049	-.0043	.0111	.0017		
(0.83)	(0.54)	(1.13)	(0.64)	(0.88)	(0.40)		
.0154	-.0035	.0021	-.0064	.0054	.0015	.0092	.0157
(1.12)	(0.61)	(0.64)	(0.93)	(0.62)	(0.37)	(0.38)	(0.88)

The table shows results from Fama-MacBeth regressions for monthly stock returns (top panel) and estimated model return residuals (bottom panel), for different specifications of independent variables. R6 is the six-month trailing return. SEO, BBK, and EFM are dummy variables which indicate that the announcement of the relevent corporate action (secondary equity offering, share buy-back, or equity-financed merger) occurred in the preceding six months. E/P and SIZE are as defined in Table 4. The data are monthly observations from 1985 through 1999. Standard errors are adjusted for induced six-month autocorrelation. The resulting t statisitics are in parentheses.

Table 9: *Dividend Initiations and Omissions*

R6	SEO	BBK	EFM	DDIV	INIT	OMIT
Dependent Variable=6-Month Returns						
.0532				.0409		
(4.93)				(6.53)		
.0530					.0315	-.0484
(4.92)					(4.01)	(4.48)
.0539	-.0177	.0079	-.0201	.0409		
(5.06)	(3.11)	1.34)	(3.47)	(6.65)		
.0537	-.0178	.0078	-.0203		.0312	-.0492
(5.04)	(3.14)	1.32)	(3.49)		(4.26)	(4.55)
Dependent Variable=Model Residuals						
.0187				.0147		
(1.07)				(1.58)		
.0189					.0209	-.0101
(1.08)					(1.72)	(0.75)
.0191	-.0020	.0048	-.0035	.0148		
(1.10)	(0.34)	1.04)	(0.50)	(1.60)		
.0193	-.0019	.0046	-.0036		.0210	-.0103
(1.11)	(0.32)	0.98)	(0.51)		(1.74)	(0.76)

The table shows results from Fama-MacBeth regressions for monthly stock returns (top panel) and estimated model return residuals (bottom panel) on lagged returns and dividend change indicators. INIT and OMIT are dummies indicating the occurrence of an initiation or omission in the six months prior to the return. DDIV is INIT-OMIT. The data are monthly observations from 1985 through 1999. Standard errors are adjusted for induced six-month autocorrelation. The resulting t statisitics are in parentheses.

Table 10: *Tests of Model Constraints*

The table shows the results of two specification tests based on the fitted model residuals. The Lagrange multiplier statistic weights the cross-sectional average product of the residuals and the instruments (See footnote 14.) The Fama-MacBeth statistic tests the hypothesis that the cross-sectional regression coefficients of the residulas on the instruments are jointly equal to zero. Both statistics are asymptotically χ^2 distributed with degrees of freedom equal to the number of instruments. Asymptotic p values are in parentheses.

Instruments:	Lagrange Multiplier	Fama-MacBeth
R6, DFY1, DLTG, SEO, BBK, EFM	7.3335 (0.29)	4.5395 (0.60)
R6, DFY1, DLTG, SEO, BBK, EFM, DDIV	11.2058 (0.13)	7.7115 (0.36)

Table 11: *Split Sample Results*

A. High Coverage

R6	SEO	BBK	EFM	DFY1	DLTG	E/P	SIZE	INIT	OMIT
Dependent Variable=6-Month Returns									
.0772	-.0132	.0090	-.0064						
(3.83)	(2.04)	(2.04)	(0.90)						
.0509	-.0152	.0088	-.0066	.0381	.0169				
(2.46)	(2.33)	(2.11)	(0.87)	(2.62)	(3.66)				
.0521	-.0164	.0060	-.0076	.0351	.0154	.0026	.0246		
(3.27)	(2.76)	(1.87)	(1.06)	(3.04)	(3.78)	(0.11)	(0.96)		
.0518	-.0161	.0058	-.0077	.0354	.0154	.0028	.0250	.0176	.0072
(3.26)	(2.72)	(1.75)	(1.08)	(3.08)	(3.76)	(0.12)	(0.99)	(1.21)	(0.41)
Dependent Variable=Model Residuals									
.0175	-.0028	.0061	-.0051						
(0.81)	(0.38)	(1.48)	(0.69)						
.0119	-.0031	.0062	-.0052	.0053	.0027				
(0.58)	(0.43)	(1.53)	(0.69)	(0.38)	(0.52)				
.0105	-.0051	.0031	-.0066	.0016	.0007	.0012	.0353		
(0.63)	(0.73)	(0.93)	(0.93)	(0.15)	(0.17)	(0.05)	(1.31)		
.0099	-.0045	.0031	-.0065	.0021	.0007	.0014	.0360	.0139	.0076
(0.60)	(0.65)	(0.88)	(0.93)	(0.19)	(0.15)	(0.06)	(1.35)	(0.77)	(0.47)

B. Low Coverage

R6	SEO	BBK	EFM	DFY1	DLTG	E/P	SIZE	INIT	OMIT
Dependent Variable=6-Month Returns									
.0867	-.0218	.0135	-.0273						
(6.74)	(2.70)	(2.21)	(3.87)						
.0663	-.0325	.0128	-.0296	.0945	.0211				
(4.51)	(4.26)	(1.83)	(2.20)	(7.55)	(3.34)				
.0719	-.0267	.0129	-.0260	.0943	.0210	.0111	-.0344		
(5.54)	(3.20)	(2.13)	(1.94)	(10.10)	(3.54)	(0.46)	(1.68)		
.0720	-.0273	.0121	-.0265	.0936	.0207	.0105	-.0343	.0218	-.0029
(5.59)	(3.32)	(1.97)	(2.00)	(9.96)	(3.49)	(0.43)	(1.69)	(1.72)	(0.14)
Dependent Variable=Model Residuals									
.0125	-.0109	.0152	-.0048						
(0.70)	(0.66)	(1.87)	(0.26)						
.0047	-.0140	.0138	-.0079	.0237	-.0018				
(0.25)	(0.83)	(1.73)	(0.43)	(1.62)	(0.19)				
.0161	-.0014	.0115	-.0078	.0114	-.0004	.0446	-.0287		
(1.02)	(0.07)	(1.51)	(0.43)	(0.90)	(0.04)	(1.92)	(1.19)		
.0142	-.0004	.0116	-.0095	.0074	-.0001	.0426	-.0276	.0321	-.0263
(0.92)	(0.02)	(1.51)	(0.52)	(0.60)	(0.45)	(1.79)	(1.11)	(1.93)	(0.79)

Appendix

A Derivation of Model Equations

This appendix derives the solution to the most general specification of our model. As a first step, we treat the simple case which is used in Section 3, which allows only one lag of forecast revisions to enter the Δi equation.

Throughout we assume returns are generated by equation (2),

$$r_t = \alpha_0 + \alpha_1' \, \xi_{t-1} + \epsilon_t^{(0)}$$

where ξ is cumulative unexpected revision defined in equation (1).

Lemma A.1 *Suppose analysts' revisions obey*

$$\Delta i_t = \mu + A \, \Delta i_{t-1} + a(L) \, r_{t-1} + \epsilon_t^{(1)}$$

which is assumed strictly stationary. Then

$$\xi_t = (I + EL)^{-1} \, F \, \lambda_t$$

where

$$\begin{aligned} F &= [I - A - a(1) \, \alpha_1']^{-1} \\ E &= F \; a(1) \; \alpha_1' \\ \lambda_t &= (\epsilon_t^{(1)} + a(1) \, r_t). \end{aligned}$$

Proof. The conclusion of the lemma is equivalent (after rearranging terms) to

$$\xi_t - A \, \xi_t = \epsilon_t^{(1)} + a(1) \, r_t + a(1) \, \alpha_1' \, \xi_t - a(1) \, \alpha_1' \, \xi_{t-1} \tag{11}$$

or

(12) $$\xi_t = (I - A)^{-1} \left[\epsilon_t^{(1)} + a(1)\left(r_t - \alpha_1'\left(\xi_{t-1} - \xi_t\right)\right)\right].$$

We establish the latter equation by simply expanding out the terms in the definition of ξ:

$$\xi_t \equiv \sum_{k=0}^{\infty} \left[\mathrm{E_t}[\Delta i_{t+k}] - \mathrm{E_{t\text{-}1}}[\Delta i_{t+k}]\right].$$

It is clear from this expression that the intercept μ does not affect ξ. So we set $\mu = 0$ below without loss of generality.

Now:

$$\begin{aligned}
E_t(\Delta i_t) &= \Delta i_t \\
E_t(\Delta i_{t+1}) &= A\,\Delta i_t + a(L) r_t \\
E_t(\Delta i_{t+2}) &= A\,E_t(\Delta i_{t+1}) + a(L) E_t(r_{t+1}) \\
&= A\,E_t(\Delta i_{t+1}) + a_1(\alpha_0 + \alpha_1'\xi_t) + a_2 r_t + ... \\
E_t(\Delta i_{t+3}) &= A\,E_t(\Delta i_{t+2}) + a_2(\alpha_0 + \alpha_1'\xi_t) + a_3 r_t + ...
\end{aligned}$$

and so on until

$$E_t(\Delta i_{t+k}) = A\,E_t(\Delta i_{t+k-1})$$

for k greater than $p + 3$ where p is the order of $a(L)$.

Defining $M_t = E_t\left[\sum_{k=0}^{\infty} \Delta i_{t+k}\right]$, and collecting terms, we then have

$$M_t = (1 - A)^{-1}\left[\Delta i_t + a(1)(\alpha_0 + \alpha_1'\xi_t) + a(1)\,r_t + \hat{a}_1\,r_{t-1} + \hat{a}_2\,r_{t-2} \ldots + \hat{a}_p\,r_{t-p}\right]$$

where $\hat{a}_k \equiv \sum_{j=k}^{p} a_j$.

Since $\xi_t = M_t - M_{t-1} + \Delta i_{t-1}$, it follows that:

$$\begin{aligned}
\xi_t &= (1 - A)^{-1}\left[\Delta i_t - \Delta i_{t-1} + a(1)\,\alpha_1'(\xi_t - \xi_{t-1}) + a(1)\,r_t - a_0\,r_{t-1} - a_1\,r_{t-2} \ldots - a_p\,r_{t-p-1}\right] + \Delta i_{t-1} \\
&= (1 - A)^{-1}\left[\epsilon_t^{(1)} + a(1) r_t + a(1)\,\alpha_1'(1 - L)\xi_t\right].
\end{aligned}$$

which is the same as (12).

QED

We now handle the more general case, which allows for two types of exogenous variables to enter into the Δi equation. In the text, these are used to model the dummy variables which code for the occurrence of corporate actions. These are modelled as either an unpredictable martingale difference (e.g. for mergers, buybacks, and SEOs) or as the difference in a stationary first-order Markov process (e.g. changes in dividend-paying status). We denote the former by Δx and the latter by ΔS.

Generalizing the notation, the full specification may then be written

$$\begin{aligned}
\Delta i_t &= \mu^{(1)} + A^{(1)}(L)\,\Delta i_{t-1} + B^{(1)}(L)\,S_{t-1} + C^{(1)}(L)\,\Delta x_{t-1} + a^{(1)}(L)\,r_{t-1} + \epsilon_t^{(1)} \\
S_t &= \mu^{(2)} + A^{(2)}(L)\,\Delta i_{t-1} + B^{(2)}(L)\,S_{t-1} + C^{(2)}(L)\,\Delta x_{t-1} + a^{(2)}(L)\,r_{t-1} + \epsilon_t^{(2)} \\
\Delta x_t &= \mu^{(3)} + \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \epsilon_t^{(3)}.
\end{aligned}$$

We are now allowing a more general lag structure in the first equation. In the text, we do not have the Δx variables affecting S. Hence we put $C^{(2)}(L) = 0$. We also have only one lag of S in the S equation. These last two restrictions are not used in the derivation below.

More important is the restriction that only ΔS, not S itself, enters the Δi equation. Hence the matrices $B^{(1)}(L)$ are of the form:

$$B_0^{(1)} = b_0,\ B_1^{(1)} = b_1 - b_0,\ \ldots B_{p+1}^{(1)} = -b_p$$

where p is the order of $B^{(1)}(L)$. The key implication of this restriction is

$$\sum_{i=1}^{p} B_i^{(1)} = 0. \tag{13}$$

To start, stack the variables in this specification, $\Delta z' \equiv [\Delta i' \;\; S' \;\; \Delta x']'$, and write the full system as

$$\Delta z_t = \mu + H(L)\,\Delta z_{t-1} + h(L)\,r_{t-1} + \epsilon_t. \tag{14}$$

Then we can define a new variable, η, in analogy with equation (1), as

$$\eta_t \equiv \sum_{k=0}^{\infty} \left[\mathrm{E_t}[\Delta z_{t+k}] - \mathrm{E_{t-1}}[\Delta z_{t+k}] \right].$$

Then η consists of three stacked components, the first of which is ξ. Label the other two, which are the long-run innovations in S and Δx, as $\eta^{(2)}$ and $\eta^{(3)}$ respectively.

Our return equation then becomes

$$r_t = \alpha_0 + \hat{\alpha}_1' \, \eta_{t-1} + \epsilon_t^{(0)}$$

subject to the restriction

$$\hat{\alpha}_1' = [\alpha_1' \;\; 0 \;\; 0]. \tag{15}$$

We can now state the result we are after.

Proposition A.1 *Assume the system given by equation (14) is strictly stationary. Then*

$$\xi_t = (I + EL)^{-1} \, F \, \lambda_t$$

where

$$\begin{aligned} F &= [I - A^{(1)}(1) - a^{(1)}(1)\, \alpha_1']^{-1} \\ E &= F \; a^{(1)}(1) \; \alpha_1' \\ \lambda_t &= [\epsilon_t^{(1)} + a^{(1)}(1)\, r_t + C^{(1)}(1)\, (\Delta x_t - \mathrm{E}[\Delta x_t])]. \end{aligned}$$

Proof. We prove this by putting the Δz system in companion form and then applying the lemma above. So suppose q is the maximal order of the lag polynomial $H(L)$. Then let

$$J = \begin{bmatrix} H_0 & H_1 & \dots & H_q \\ I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & & \ddots & \vdots \end{bmatrix}$$

and

$$y_t' = [\Delta z_t' \ \ \Delta z_{t-1}' \ \dots \ \Delta z_{t-q}']'$$

and write equation (14) as

$$y_t = \bar{\mu} + J\, y_{t-1} + \bar{h}(L) r_{t-1} + \bar{\epsilon}_t \tag{16}$$

where $\bar{\mu}, \bar{h}(L)$, and $\bar{\epsilon}$ are vectors whose first components are $\mu, h(L)$, and ϵ and whose other components are zeros.

Then, again, in analogy with ξ and η, we define ζ as

$$\zeta_t \equiv \sum_{k=0}^{\infty} [\mathrm{E_t}[y_{t+k}] - \mathrm{E_{t\text{-}1}}[y_{t+k}]]$$

so that the returns are given by

$$r_t = \alpha_0 + \bar{\alpha}_1'\, \zeta_{t-1} + \epsilon_t^{(0)}$$

with $\bar{\alpha}_1$ restricted as in (15), $\bar{\alpha}_1' = [\hat{\alpha}_1' \ \ 0 \ \ 0 \dots 0]$.

Now we can apply our lemma to the stationary AR(1) system (16) and deduce

$$\zeta_t = (I - J)^{-1} \ \left[\bar{\epsilon}_t + \bar{h}(1)\left(r_t - \bar{\alpha}_1'\left(\zeta_{t-1} - \zeta_t\right)\right)\right] \tag{17}$$

which follows from just rearranging the lemma's conclusion.

It remains to unwind this expression and recover first η and then ξ in terms of the original quantities in the specification. So first note that η_t is the first component of ζ_t, which means we only need analyze the first row of the matrix $(I - J)^{-1}$. Also the quantity in square brackets is

$$\begin{bmatrix} [\epsilon_t + h(1)\left(r_t - \hat{\alpha}_1'\left(\eta_{t-1} - \eta_t\right)\right)] \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

Next, it is easy to show that the (1,1) element of $(I - J)^{-1}$ is just $(I - H(1))^{-1}$. Hence we have shown that (17) implies

$$\eta_t = (I - H(1))^{-1} \; \left[\epsilon_t + h(1) \left(r_t - \hat{\alpha}_1' \left(\eta_{t-1} - \eta_t\right)\right)\right]. \tag{18}$$

Multiply this equation by $(I - H(1))$. Since, using the restriction (13), $H(1)$ is

$$\begin{bmatrix} A^{(1)}(1) & 0 & C^{(1)}(1) \\ A^{(2)}(1) & B^{(2)}(1) & C^{(2)}(1) \\ 0 & 0 & 0 \end{bmatrix}$$

and $h(1)\,\hat{\alpha}_1'$ is

$$\begin{bmatrix} a^{(1)}(1)\,\alpha_1' & 0 & 0 \\ a^{(2)}(1)\,\alpha_1' & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

the first line of the resulting formula says

$$\xi_t - A^{(1)}(1)\,\xi_t - C^{(1)}(1)\,\eta_t^{(3)} = \epsilon_t^{(1)} + a^{(1)}(1)\,r_t + a^{(1)}(1)\,\alpha_1'\,\xi_t - a^{(1)}(1)\,\alpha_1'\,\xi_{t-1}. \tag{19}$$

Also, the third line of the same formula is simply $\eta_t^{(3)} = \epsilon_t^{(3)}$ which is just Δx_t with the mean removed. Plugging this in to the left side of (19) and rearranging terms yields the desired conclusion.

QED

Note that, for consistency with the previous section, equation(10) in the text drops the superscripts of the coefficient matrices in utilizing the conclusion of the proposition.

Also note that neither the innovations to the stationary process S, nor any of its coefficient matrices B, enter the resulting formula for ξ. This property is used in Section 4.

B Description of Data and Variable Definitions

IBES Normalizations

IBES monthly consensus forecast record fields are transformed into the two variables DFY1 and DLTG used inthe text as follows.

For DFY1,

$$\text{DFY1}(t) = k(n)\frac{\text{FYE}(t) - \text{FYE}(t-6)}{P(t-6)}$$

where t indexes the month where the forecasts are made, $\text{FYE}(t)$ is the mean earnings estimate for the next full fiscal year following the one that was underway at $t-6$. $P(t)$ is the stock price at month t, and $k(n)$ is a normalising constant to ensure that the forecast changes at t are comparable across firms with different fiscal years. Specifically, if the "next" fiscal year is already underway at $t-6$, then k is the reciprocal of the amount of that year remaining, up to a maximum value of 3.

For DLTG,

$$\text{DLTG}(t) = \text{LTG}(t) - \text{LTG}(t-6)$$

where $\text{LTG}(t)$ is the mean long term growth rate estimate at time t.

Both DFY1 and DLTG are expressed as a percentile rank relative to the other firms existing in the cross section at time t. These are then centered at zero by subtracting the median, to yield a number between -1/2 and 1/2.

SDC Event Definitions

Secondary Equity Offerings: These are obtained from the SDC Global New Issues database from 1978 to 1999, searching for all public & private issues of common stock by a US company where the issue is not an IPO. Announcement dates were taken from the D (issue date) data field. This is defined as the date when the securities were offered.

Buybacks: These are obtained from the SDC domestic mergers database for all US stocks from 1978 to 1999. Included are all deals in which the company buys back its equity securities or securities convertible into equity, either on the open market, through privately negotiated transactions, or through a tender offer. Announcement dates were taken from the DA (date announced) data field.

Equity Financed Mergers: These are obtained from the SDC domestic mergers database for all US stocks from 1978 to 1999 where the form of the deal is a merger or acquisition. Only completed deals are included in the database. All events where the payment form contained no cash were used in the sample. The Announcement date is defined as the date when either target or acquiror makes a public announcement that it held negotiations, or received a formal proposal to combine, acquire, recapitalize, etc.

References

Abarbanell, Jeffery, 1991, Do analysts earnings forecasts incorporate information in prior stock price changes?, *Journal of Accounting and Economics* 14, 147–165.

Abel, Andrew B., and Olivier Blanchard, 1986, The present value of profits and cyclical movements in investment, *Econometrica* 54, 249–274.

Amemiya, Takeshi, 1985, *Advanced Econometrics* (Harvard University Press: Cambridge).

Asquith, Paul, and David W. Mullins, 1986, Equity issues and offering dilution, *Journal of Financial Economics* 15, 61–89.

Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *Journal of Finance* 51, 1681–1714.

Chan, Wesley, 2001, Stock price reaction to news and no-news: Drift and reversal after headlines, M.I.T. Working Paper.

Dechow, Patricia, Amy Hutton, and Richard Sloan, 1999, The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings, University of Michigan Working Paper.

Falk, Barry, and Bong-Soo Lee, 1998, The dynamic effects of permanent and transitory labor income on consumption, *Journal of Monetary Economics* 41, 371–387.

Fama, Eugene F., 1998, Market efficiency, long-term returns and behavioral finance, *Journal of Financial Economics* 49, 283–306.

———, and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gilchrist, Simon, and Charles P. Himmelberg, 1995, Evidence on the role of cash flow for investment, *Journal of Monetary Economics* 36, 541–572.

Hong, Harrison, Terrence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181–208.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Kadiyala, Padma, and P. Raghavendra Rau, 2001, It's all under-reaction, Purdue University Working Paper.

LaPorta, Rafael, 1996, Expectations and the cross-section of expected returns, *Journal of Finance* 51, 1715–1742.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

Loughran, Tim, and Anand M. Vijh, 1997, Do long-term shareholders benefit from corporate acquisitions?, *Journal of Finance* 52, 1765–1790.

Michaely, Roni, Richard H. Thaler, and Kent Womack, 1995, Price reaction to dividend initiations and omissions: Overreaction or drift?, *Journal of Finance* 50, 573–608.

Mitchell, Mark, and Erik Stafford, 2000, Managerial performance and long-term stock price performance, *Journal of Business* 73, 287–328.

Pritamani, Mahesh, and Vijay Singal, 2001, Return predictability following large price changes and information releases, *Journal of Banking and Finance* 25, 631–656.

Quah, Danny, 1990, Permananent and transitory movements in labor income: An explanation for "excessive smoothness" in consumption, *Journal of Political Economy* 98, 449–475.

Speiss, Katherine D., and John Affleck-Graves, 1995, Underpreformance inlong-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243–268.

Teoh, Siew H., and T. J. Wong, 2002, Why do new issues and high-accrual firms underperform? The role of analysts' credulity, *Review of Financial Studies* forthcoming.

Travlos, Nickolaos G., 1987, Corporate takeover bids, method of payment, and bidding firms' stock returns, *Journal of Finance* 42, 943–963.

Vuolteenaho, Tuomo, 2002, What drives firm-level returns?, *Journal of Finance* 57, 233–264.