Andrew Jackson
Timothy Johnson
London Business School

Unifying Underreaction Anomalies*

This paper asks whether momentum and postevent drift are manifestations of the same underlying mechanism or are separate phenomena. We find that both effects can be attributed to persistence in returns following news that affects expected earnings or earnings growth. Holding these quantities fixed, there is no momentum effect, nor is there postevent drift for our sample of events, which includes seasoned equity offerings, repurchases, equity-financed mergers, and dividend initiations and omissions. The implication is that return continuation follows fundamental news in general, and in aggregate, this explains momentum.

I. Introduction

Two of the more convincingly documented stock market anomalies are momentum and postevent 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 6 months to a year.[1] Postevent 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 1 to 3 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

* We are grateful to Stefan Nagel, Allen Poteshman, Tuomo Voulteenaho, and an anonymous referee for thoughtful comments. We also thank seminar participants at the London Business School 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. Contact the corresponding author, Timothy Johnson, at tjohnson@london.edu.

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

(*Journal of Business,* 2006, vol. 79, no. 1)

0021-9398/2006/7901-0004$10.00

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 usually have 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 focused 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 postevent drift across all classes of events. If, in fact, future returns are high following any type of good news generally, then presumably a portfolio of the best recently performing stocks, which ought to contain the firms with the most good news, also will include the best subsequent performers. Mechanically, 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.

The most straightforward way, conceptually, to carry out our investigation would be to undertake a giant bookkeeping 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, Wesley Chan (2003) did something like this for a sample of 1,557 stocks. He identified all news events for each over the period 1980–99 and demonstrated both that there is a generic postevent 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 postformation period.[3]

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

These results strongly suggest that the conjecture of a single underlying effect may be correct.

The next step is to actually link the return dynamics to the news for each stock. That is, we would like to establish that the continuation of returns is not due to coincident momentum (for the firms that have news) and show that the absence of continuation (for the rest) is due to the absence of news and not the absence of momentum. Those are our goals here. We seek to build on Chan's results by providing a model that explicitly disentangles the news reaction from momentum.

Our aim is to be able to write down a tractable specification of the cross section of expected returns 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 return by reference to some omnibus event table, we may be able to find instrumental variables that signal the occurrence of whatever type of news 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 account for diverse postevent 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 postevent 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, 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, we leave it for future work to test this theory against the alternative, perhaps less plausible, view that good news tends to coincide with increases in risk. In short, we do not address the deeper question of why news that affects earnings expectations has persistent effects. But we think that identifying this association between company fundamentals and return continuation considerably sharpens the focus of the theoretical effort to understand what is going on.

Our work is closely related to two other recent papers. Louis 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 employed what we think are improved measures of expectations and claim that these can finish the job.

In a similar vein, Cohen, Gompers, and Vuolteenaho (2002) find that, holding expected cash flows constant, purely transitory price increases are not followed by further price increases; hence, there is no residual momentum at an annual frequency. They reconcile their results with Chan et al. (1996) by noting that their vecxtor autoregression (VAR) equations better capture new information in cash flows than the Standardized Unexplained Earnings (SUE) measure of Chan et al. (1996). Our innovation with respect to Cohen et al. (2002) is to extend this analysis to postevent drift anomalies using a unified consistent methodology. Additionally, we measure changes in expected earnings directly via revisions to analyst forecasts, rather than indirectly through an accounting-based present value formula.

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 III fits the basic model and demonstrates that it explains momentum. In Section IV, 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 then is shown to explain the abnormal returns subsequent to those events. Section V presents the results of various robustness checks. The final section contains some concluding remarks on the interpretation of our findings.

II. 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 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 used changes in consensus earnings forecasts (and growth forecasts) from the Institutional Brokers Estimate System (I/B/E/S), 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 the nonforecastability 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 I/B/E/S 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^* is a vector of parameters that affects a firm's value and $\{i_t\}$ is any (noisy) series that eventually converges to i^*. Then, $E_t(i^*) = i_t + E_t\left(\sum_{k=0}^{\infty} \Delta i_{t+k}\right) = i_t + \sum_{k=0}^{\infty} E_t(\Delta i_{t+k})$, (assuming the latter is finite). The correct measure of the amount of news about i^* that arrives at time t then is

$$\xi_t \equiv E_t(i^*) - E_{t-1}(i^*) = \sum_{k=0}^{\infty} [E_t(\Delta i_{t+k}) - E_{t-1}(\Delta i_{t+k})]. \qquad (1)$$

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, in fact, never converge to the true value, as long as any bias component is stationary. In our setting, we interpret the observed I/B/E/S 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

$$r_t = \alpha_0 + \alpha'_1 \xi_{t-1} + \varepsilon_t^{(0)}. \qquad (2)$$

That is, innovations in our fundamental variables cause changes in expected returns and other things do not.[4] The basic model (2) envisions only 1-period-ahead effects (so $E_t(r_{t+2})$ is a constant), which is overly restrictive. We have also fit two-lag and geometric lag models (see Section V.D).

4. 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.

The best specification, in general, will depend on the time interval Δt. The simplest version, however, suffices to illustrate our main points. At this stage, 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.[5] 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

$$\Delta i_t = \mu + A(L)\Delta i_{t-1} + a(L)\tilde{r}_{t-1} + \varepsilon_t^{(1)}. \tag{3}$$

Here μ is a constant, A and a are polynomials (of appropriate dimension) in the lag operator L, and $\tilde{r} \equiv r - \alpha_0$ is the stock return in excess of its unconditional expectation. Equation (3) is just a linear specification in terms of observable quantities. And, given its coefficients, we can compute the pure forecast innovation series ξ via (1). However, the computation also depends on the parameters in (2), because expectations of future returns are involved.

Despite this complication, the resulting expression for the true change in conditional expectations can be written compactly as

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

where

$$F = [I - A(1) - a(1)\alpha'_1]^{-1} \tag{5}$$

$$E = Fa(1)\alpha'_1 \tag{6}$$

$$\lambda_t = [\varepsilon_t^{(1)} + a(1)\tilde{r}_t]. \tag{7}$$

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

$$r_t = \alpha_0 + \alpha'_1 (I + EL)^{-1} F\lambda_{t-1} + \varepsilon_t^{(0)}. \tag{8}$$

Intuitively, equation (8) is just a regression of returns on a filtered version of the series $\{\lambda\}$. That series is the augmentation described previously of the current unexpected change in the observed forecast by the cumulative future revision predicted by current returns. A further filtering step is required (implemented by the operator $(I + EL)^{-1}$) only because of the added complication that the return part of λ_t is itself

5. To our knowledge, this was first documented by Abarbanell (1991).

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 nonlinear model for returns. This means that , ultimately, we will not be able to gauge our model by the familiar method of running ordinary least squares (OLS) regressions of returns on some new candidate variables alongside established competitors. Instead, we first estimate (3), then fit the model (8) (via nonlinear least squares), and then evaluate it by checking for remaining predictability in the unexpected component of returns.

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

First, our approach is closely analogous 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 are not directly observable.[6] This permanent component, call it X, is analogous 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, even though, of course, it is 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, then using the fitted model to solve for the expected present value of future profits and, from this, deducing the true "marginal Q," or return

6. See Quah (1990) for the development of this idea, which was implemented by Falk and Lee (1998).

on investment.[7] Again, this construction is much like our computation of ξ. The issue under study is whether the synthetic Q alone can explain firms' current investment or current cash flow also matters. But current cash flow (like past returns in our setup) is also one of the most important instruments used in the first-stage structural model. Still, we regard the null as supported if cash flow enters the investment model only via this constrained channel.

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 in describing analysts' behavior, and since Δi is fit independently of the return model, the predictors do not subsequently enter 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 explicitly test the model's restrictions by, in effect, giving the data a chance to undo the constraints.

The next section introduces the data series we employed 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, in fact, do account for momentum effects. After that, we turn to the connection to postevent drift in Section IV.

III. Initial Results

As discussed at the start of the preceding section, the initial expectational variables employed in this study are based on I/B/E/S consensus earnings forecasts. Our working hypothesis is that good news generates persistent positive abnormal returns if and only if the I/B/E/S forecasts, corrected as described previously, concurrently increase, with negative abnormal returns likewise when they decrease.

One could imagine different types of news affecting earnings at different horizons and, hence, distinct informational roles for changes in long-term and short-term forecasts. For each stock tracked by I/B/E/S, the monthly data records include average forecasts for a number of fiscal horizons, as well as a separate consensus forecast for the long-term (5-year) growth rate. We employ both revisions to the nearest fiscal year's estimate (which we call DFY1) and revisions to the long-term growth number (called DLTG) as measures of shocks to fundamentals.

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

The intuition of a simple Gordon-type model of stock prices suggests that these two should capture the most relevant characteristics of firms' expected cash-flow profiles.

Our short-term earnings measure parallels that used in Chan et al. (1996) and many other studies. I/B/E/S long-term growth rates have been studied by LaPorta (1996), who documents an independent informational role for them in predicting stock returns. The forecasting properties of this variable recently were examined extensively by Chan, Karceski, and Lakonishok (2003). Those authors document that actual earnings growth rates are very difficult to forecast and predictive regressions achieve low values of R^2. The I/B/E/S variable, however, does appear to have information content, entering significantly into regressions of 1- to 3-year growth of sales and operating income before depreciation.

Of course, the fact that other variables also enter into the regressions of Chan et al. (2003) immediately implies that analyst forecasts are not true expectations: they neglect other meaningful predictors. Some of the reasons for this are simply mechanical (e.g., delayed revisions by analysts), while others may have to do with incentives and institutions. These failings are the reason why the first step in our analysis must be to correct the I/B/E/S revisions for their most obvious weaknesses. If we fail to do so completely or the underlying information content of the numbers was negligible to begin with, this will only make it less likely that our corrected proxies will succeed in accounting for the predictability in stock returns.

The basic time unit of the study is 6-month intervals. For each stock, our fundamental variable Δi_t comprises the two components DFY1 and DLTG, defined respectively as the change over the preceding 6 months in the mean forecast for the next fiscal year and the change in the mean long-term growth forecast over the same period.[8]

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 et al. (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 stage of our program is to fit the prediction equation (3) for our bivariate process Δi to extract its innovation component. Table 2 shows estimates of a specification which incorporates one lag of Δi itself and two lags of returns.[9] 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.

8. A complete description of all data fields and variables is given in Appendix B.
9. Including longer lags has no significant effects on the results that follow.

TABLE 1 Stocks with I/B/E/S Data

Year	Number of Firms	Average Size Decile: Mean	Average Size Decile: Median	Stocks by Exchange: NYSE	Stocks by Exchange: AMEX	Stocks by Exchange: NASDAQ
1983	1,546	4.1	3.0	1,140	107	299
1984	1,715	3.4	2.0	1,140	119	456
1985	1,945	3.3	1.0	1,141	137	666
1986	1,963	3.3	2.0	1,139	135	688
1987	2,007	3.2	2.0	1,113	144	749
1988	1,982	2.7	1.0	1,005	139	838
1989	1,979	2.8	1.0	1,026	142	811
1990	2,058	2.8	1.0	1,020	157	881
1991	2,027	2.9	2.0	1,009	133	885
1992	2,050	3.2	2.0	1,070	120	859
1993	2,181	3.4	2.0	1,136	114	930
1994	2,452	3.6	2.0	1,227	97	1,126
1995	2,638	3.8	3.0	1,278	82	1,275
1996	2,791	4.0	3.0	1,362	74	1,353
1997	3,069	4.0	3.0	1,456	80	1,531
1998	3,355	4.2	3.0	1,568	89	1,697
1999	3,435	4.6	4.0	1,621	101	1,712

NOTE.—The table describes the stocks for which there are I/B/E/S forecasts for both the next fiscal year's earnings and long-term (5-year) growth. For a firm to be included in a given month, it must have forecasts for at least the previous 6 months and stock prices on CRSP for the previous 12 months and the following 6 months. Size deciles are computed using month-end NYSE breakpoints.

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, therefore, the significance of lagged revisions is not surprising. The finding that near-term forecasts underreact to stock returns at lags up to 1 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 dependency, analysts actually overadjust 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). In addition to having quite different dynamics, the

TABLE 2 Forecasting Forecast Revisions

	μ	$\mathrm{DFY1}_{t-1}$	DLTG_{t-1}	R6	LR6
$\mathrm{DFY1}_t$	.0021	.2718	−.0012	.3219	.0924
	(1.60)	(38.03)	(.43)	(20.51)	(9.95)
DLTG_t	.0006	−.0071	−.0573	.2052	.0744
	(.60)	(1.21)	(11.07)	(23.02)	(9.45)

NOTE.—The table shows estimates for the semiannual specification: $\Delta i_t = \mu + A\Delta i_{t-1} + a_1 r_{t-1} + a_2 r_{t-2} + \varepsilon_t^{(2)}$, where Δi is the bivariate process of I/B/E/S forecast revisions whose components are DFY1 and DLTG, the percentile rank changes in next-fiscal-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 6-month autocorrelation. The resulting t statistics are in parentheses.

TABLE 3 News Premia Estimates

	α_1	(t Stat)	Average SD CS	Average SD TS
ξ_1	.0514	(3.03)	.351	.377
ξ_2	.0159	(2.15)	.237	.246

NOTE.—The first column of the table gives return premium estimates, α_1, in the return model $r_t = \alpha_0 + \alpha_1' \xi_{t-1} + \varepsilon_t^{(1)} \xi_{t-1}$, where ξ is the corrected news process defined by equations (4)–(8). The parameters are estimated by nonlinear least squares (see note 11). Estimates of the ξ process, which depend on α_1, then are 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.

two components of Δi also clearly are responding to different types of fundamental news. On average, the estimated residuals of DFY1 and DLTG for a given firm have opposite signs 47% of the time and average correlation of only 0.12.

The next step is to estimate 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.[10] Given these values for A and a, and the fitted Δi residuals $\hat{\varepsilon}^{(2)}$, the only unknowns in equation (8) are the values of α. We fit these by nonlinear least squares (NLLS) with a time-varying intercept. The resulting pair of return premia, the two components of α_1, is shown in table 3.[11]

Innovations to near-term expected earnings have about three times the effect that innovations to long-run growth do on expected returns. For reference, the table also shows the average cross-sectional and time-series variability[12] of the two components of the innovation series. Combined with the α estimates, these indicate that the magnitude of the fitted model's effects are on the order of ± 250 basis points of return differential (in expected 6-month returns), again, with 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 6-month returns) as a predictor. The first panel shows results when the raw returns themselves (over the next 6 months) are the dependent variable. Here we confirm that the momentum anomaly is strongly present in our sample of returns. We replicate the finding

10. By not accounting for the estimation uncertainty in these, we overstate the accuracy of our estimate of the return parameter α_1 in the following. On the other hand, when we test the unexpected component of returns for remaining underreaction effects, the t-statistics overstate significance levels; i.e., we bias the tests in the direction of rejecting our model.

11. Standard errors are computed from the asymptotic covariance matrix $T^{-1}A^{-1}BA^{-1}$, 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.

12. These are arithmetic averages of standard deviations. Root mean squares were similar. The average time-series correlation between the two components is 0.35.

TABLE 4 Return Regressions

R6	DFY1	DLTG	E/P	SIZE
Dependent Variable = Raw 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)	(.61)
Dependent Variable = Unexpected Returns				
.0178				
(1.07)				
.0127	.0088	.0007		
(.80)	(.75)	(.18)		
.0155	−.0029	.0019	.0264	.0148
(1.20)	(.14)	(.51)	(1.16)	(.91)

NOTE.—The table shows results from monthly Fama-MacBeth regressions for 6-month raw returns (top panel) and estimated unexpected returns (bottom panel), for different specifications of independent variables. R6, DFY1, and DLTG are the 6-month trailing returns and I/B/E/S 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 6-month autocorrelation. The resulting t statistics are in parentheses.

of Chan et al. (1996) that revisions to the near-term forecast neither subsume nor are subsumed by past returns as a predictor of future returns. We also document a further, independent predictive role for changes in long-term growth forecasts, supporting the view that these changes convey additional information beyond that in current-year revisions.

In the second panel, the dependent variable is instead the unexpected component of returns in excess of that fitted by our model. The message is clear. The initial predictability has been almost entirely erased by the model. This is the paper's first main result.[13]

To be careful, the conclusion here 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 as proxies for future revisions to earnings forecasts.

It is important to point out that, although past returns indeed are 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 I/B/E/S 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 then would take the opportunity

13. In Section V, we subject this finding to a variety of robustness checks, including the formation of anomaly portfolios and split sample regressions.

TABLE 5 Tests of Model Constraints

Null	Instruments	Lagrange Multiplier	Fama-MacBeth
Model	R6, DFY1, DLTG	1.5085 (.68)	1.6336 (.65)
Alternative 1	R6, DFY1, DLTG		17.1222 (.00)
Alternative 2	R6, DFY1, DLTG		29.4490 (.00)
Alternative 3	R6, DFY1, DLTG		17.2048 (.00)

Note.—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 note 14.) The Fama-MacBeth statistic tests the hypothesis that the cross-sectional regression coefficients of the residuals on the instruments are jointly equal to zero. Both tests are asymptotically distributed as $\chi^2(3)$. Alternative restrictions tested are that ξ_1 = R6 and ξ_2 = DFYI(1); ξ_1 = R6 and ξ_2 = DLTG(2); and ξ_1 DFYI and ξ_2 = DLTG(3). Asymptotic p values for all tests are in parentheses.

to undo the spurious inclusion of the I/B/E/S 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 by the analyst revision equation. With that in mind, we tested whether loosening these restrictions significantly improves the fit of the model. This is carried out via a Lagrange multiplier-type test statistic. Specifically, the test envisions the alternative hypothesis as allowing the coefficients on R6, DFY1, and DLTG to be unrestricted, and measures the incremental explanatory power of perturbing them in the direction they would want to go.[14]

A similar test may be based on the coefficients in the cross-sectional regressions of the residuals on the instruments. Here, we can construct another χ^2 statistic to test the restriction that these are jointly zero. This is analogous to the usual F statistic, corrected for the induced autocorrelation and cross correlation of the regression coefficients. We label this the Fama-MacBeth statistic in the table.

Neither test is able to reject the null hypothesis that our constraints hold. Could arbitrary combinations of the predictors, in place of our ξ do as well? The table checks the Fama-MacBeth test's ability to discriminate

14. The statistic is computed as $g_r \hat{W} g_T'$, where $g_T = 1/T \Sigma h_t, h_t = 1/N \Sigma_n \hat{u}_{n,t} X_{n,t}$, $\hat{u}$ are the residuals, $X_{n,t}$ is the vector of derivatives of the restricted specification for the nth asset with respect to the constrained variables' coefficients, and $\hat{W}$ is an estimator of the asymptotic covariance matrix of the h process. This may also be viewed as a generalized method of moments test of orthogonality in cross-sectional mean of instruments and residuals. Under the null hypothesis, the statistic is asymptotically distributed as $\chi^2(3)$ and is equivalent asymptotically to the sample size times the R^2 from a pooled regression of the model residuals on the instruments, corrected for overlapping observations.

against three natural alternatives. The first (alternative 1) is that ξ_1 is R6 and ξ_2 is DFY1, so that expected returns are just the best linear combination of these two, leaving out DLTG. Likewise, alternatives 2 and 3 are the best linear models leaving out R6 and DFY1, respectively. All these are easily rejected by the test: the residuals retain the predictable patterns in the raw data. There is no reason to think that other arbitrary rearrangements of our instruments would fare any better. If the restrictions imposed by our model in the first stage were wrong, the tests would expose them.

In the next section, we extend the results here 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.

IV. Postevent Drift

Our inferential approach to postevent 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 that is 1 for firm i at time t if the announcement date of such an event occurred for firm i in the 6 months prior to (and including) t. If the dependent variable in the regression is the return in the following 6 months, then we are implicitly examining the average postevent drifts for months $+1$ to $+12$ in event time. This setup 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 postevent 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 I/B/E/S 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 3-day announcement returns is the same

15. Some of the primary work on these can be found in Asquith and Mullins (1986), Loughran and Ritter (1995), and Spiess 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.

TABLE 6 Corporate Events in I/B/E/S Sample

	Number of Events		
Year	SEO	BBK	EFM
1983	1,156	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	1,202	550	329
1994	836	950	411
1995	1,007	999	479
1996	1,335	1,335	544
1997	1,118	1,195	662
1998	758	1,739	631
1999	850	1,363	533
Mean event return	−.0069	.0186	.0030
Mean 1-year return	−.0432	.0307	−.0392

NOTE.—The table describes the corporate action sample used in Section IV. Announcement dates for secondary equity offerings (SEOs), share buybacks (BBK), and equity-financed mergers (EFM) are taken from SDC for all firms for which we have the I/B/E/S fields used in Section III. Also shown are the mean 3-day announcement return across the whole sample and the mean 12-month return commencing in the first postannouncement month. Announcement returns are excesses over the S & P 500 return for the same period. One-year returns are relative to characteristic-matched benchmark portfolios using the technique of Daniel et al. (1997).

as that of returns for the following 12 months (starting with the first postannouncement 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 have not identified the real event date correctly in a lot of cases.[17] In any case, the 1-year characteristic-adjusted drifts are all economically large and in the same direction documented in previous research.

Interestingly, 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 misspecified (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 undo the underreaction of the analysts.

17. We do not use the announcement returns for anything that follows, so this is immaterial.

TABLE 7 Corporate Events as Predictors of Forecast Revisions

	Event Type		
Coefficient sum	SEO	BBK	EFM
Dependent = DFY1	−.0958	.0699	.0000
Dependent = DLTG	−.2 157	.0574	.0000
Lags selected	11	9	0
χ^2 drop lag	.0440	.0085	—
χ^2 add lag	.0785	.3921	.4470

NOTE.—The table shows the sum of the estimated coefficients on lags of corporate event indicator variables (cf, 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 buybacks (BBK), and equity-financed mergers (EFM).

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

$$\Delta i_t = \mu + A(L)\Delta i_{t-1} + a(L)\tilde{r}_{t-1} + c(L)\Delta x_{t-1} + \varepsilon_t^{(1)}, \quad (9)$$

where now Δx_i is the event indicator process. (For the exposition, we assume there is only one type of event. The version we fit there will have separate process and a separate lag polynomial for each of our three types.) Next, assume x_t is a martingale. This also is for simplicity, but it is not far off.[18] Then, as shown in Appendix A, equations (4), (5), and (6) are formally unchanged. But now (7) becomes

$$\lambda_t = \varepsilon_t^{(1)} + a(1)\tilde{r}_t + c(1)(\Delta x_t - E[\Delta x_t]). \quad (10)$$

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 detect no significant analyst underreaction. But for buybacks and SEOs, we find significant coefficients, nearly all of the same sign, out to lags of nearly 6 years for revisions to both year-1 estimates and long-term growth forecasts. This finding, although ancillary to the main purpose of our study, is nevertheless remarkable in its own right. We have documented a long-horizon postevent drift in analysts' forecasts that parallels closely the pattern in returns.[19]

Continuing now with our program, incorporating the lags of the event indicators in our Δi specification, we use the new estimates of the

18. In fact, previous buybacks tend to predict buybacks and momentum predicts SEOs. But the effects are weak. Correcting for this slightly strengthens our results.

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

TABLE 8 Return Regressions

R6	SEO	BBK	EFM	DFY1	DLTG	E/P	SIZE
Dependent Variable = Raw 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)	(.23)	(.55)
Dependent Variable = Unexpected Returns							
.0200	−.0016	.0053	−.0038				
(1.14)	(.26)	(1.18)	(.54)				
.0140	−.0033	.0049	−.0043	.0111	.0017		
(.83)	(.54)	(1.13)	(.64)	(.88)	(.40)		
.0154	−.0035	.0021	−.0064	.0054	.0015	.0092	.0157
(1.12)	(.61)	(.64)	(.93)	(.62)	(.37)	(.38)	(.88)

NOTE.—The table shows results from monthly Fama-MacBeth regressions for 6-month raw returns (top panel) and estimated unexpected returns (bottom panel), for different specifications of independent variables. R6 is the 6-month trailing return. SEO, BBK, and EFM are dummy variables that indicate that the announcement of the relevant corporate action (secondary equity offering, share buyback, or equity-financed merger) occurred in the preceding 6 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 6-month autocorrelation. The resulting t statistics are in parentheses.

parameters in (9) to again construct the innovation series λ_t.[20] As before, we then use λ_t to simultaneously estimate the return response coefficients α and build the corrected news process ξ_t and, from these, deduce the model's expected returns for each stock for each month. The unexpected component of realized returns 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 postevent drift. Interestingly, almost all the postmerger drift is accounted for by the raw I/B/E/S 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 indeed isolates the true informational component of the forecast series.

The second panel of the table runs the same regressions with the unexpected returns. All the event dummies—as well as momentum itself again—are now reduced to levels indistinguishable from zero.[21] Once

20. 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 III.

21. We have also used a design where the dummy variable in the regressions is set equal to the sign of the (3-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.

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 postevent drift but rather that there is as much as would be expected if the only role of the event variables were 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, although 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 previously.

Stochastically, there is an important difference between the event indicators for dividend changes and those for the corporate actions used earlier. 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 1 if a firm is a current dividend payer at time *t,* and 0 if it is not, then S is a first-order Markov process. Its transitions, ΔS, are $+1$ for initiations and -1 for omissions, which we model as

$$\Delta S_t = \mu^{(2)} + A^{(2)}(L)\Delta i_{t-1} + B^{(2)}(L)S_{t-1} + a^{(2)}(L)\tilde{r}_{t-1} + \varepsilon_t^{(2)}$$

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.

TABLE 9 Dividend Initiations and Omissions

R6	SEO	BBK	EFM	DDIV	INIT	OMIT
Dependent Variable = Raw 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 = Unexpected Returns						
.0187				.0147		
(1.07)				(1.58)		
.0189					.0209	−.0101
(1.08)					(1.72)	(.75)
.0191	−.0020	.0048	−.0035	.0148		
(1.10)	(.34)	(1.04)	(.50)	(1.60)		
.0193	−.0019	.0046	−.0036		−.0210	−.0103
(1.11)	(.32)	(.98)	(.51)		(1.74)	(.76)

Note.—The table shows results from monthly Fama-MacBeth regressions for 6-month raw returns (top panel) and estimated unexpected returns (bottom panel) on lagged returns and dividend change indicators. INIT and OMIT are dummies indicating the occurrence of an initiation or omission in the 6 months prior to the return. DDIV is INIT − OMIT. The data are monthly observations from 1985 through 1999. Standard errors are adjusted for induced 6-month autocorrelation. The resulting t statistics are in parentheses.

We have to rebuild our model's expected returns, 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 and unexpected returns 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 et al. (1995). This is important because our sample is mostly disjoint from theirs (which ends in 1988). In our data, too, there is indeed strong postevent 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

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

TABLE 10 Tests of Model Constraints

Instruments	Lagrange Multiplier	Fama-MacBeth
R6, DFY1, DLTG, SEO, BBK, EFM	7.3335 (.29)	4.5395 (.60)
R6, DFY1, DLTG, SEO, BBK, EFM, DDIV	11.2058 (.13)	7.7115 (.36)

NOTE.—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 note 14.) The Fama-MacBeth statistic tests the hypothesis that the cross-sectional regression coefficients of the residuals 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.

the components of our constructed ξ variable. Since none of the tests rejects at conventional levels, there seems little reason to fear that the forecast revision model has somehow merely rearranged the predictor variables in an arbitrary fashion. We also note that these tests are more powerful than the ones in Section III, due to the added instruments. Intuitively, it becomes increasingly implausible that our return model, with only two free parameters, simultaneously could 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 impose highly stringent conditions on the return specification.

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 therefore is 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.

V. Further Checks

To check that the results of the last section are not artifacts of the sample or methodology, we now subject them to a variety of robustness tests.

A. Split Samples

First, to see if the model's explanatory power is 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 often are 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) 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. Although it is not obvious why this kind of misspecification would make it easier for us to explain the underreaction anomalies, splitting the sample by coverage seems like a sensible check.

Table 11 shows our before and after return regressions for both the high-coverage and the low-coverage subsamples. 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 remains in these firms. The model is no less successful with them, however. Unexplained returns again have 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.

We can similarly check whether the model's explanatory power is consistent across time. In table 12, we repeat the regressions from table 4 for the first and second halves of the sample period.[24] Importantly, we do not allow ourselves to reestimate the risk premia α_1 in the two subsamples, nor do we separately estimate the innovation variables ξ. Fixing these at their full-sample estimated values increases the likelihood of detecting misspecification.

No such misspecification is apparent, however. Instead, the model eliminates momentum in both periods. The main difference is the significant presence of the SIZE and E/P control variables in the first half. But these have no effect on the inferences at issue.

B. Industry Momentum

A third way of challenging the model is to ask whether it can also explain other characteristics of momentum profits. One prominent characteristic is the strengthening of momentum in industry portfolios documented by Moskowitz and Grinblatt (1999). We have not taken industry effects into account anywhere in fitting our specification. Do the model's unexpected returns still contain such an industry component?

In table 13, we split the independent variable R6 into an industry return component (using the 20-sector classification of Moskowitz and Grinblatt 1999) and a firm-specific component, that is, net of the industry return.[25] The top panel confirms the presence of an enormous

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.

24. This table drops the event variables to be able to use the full sample period. The results are the same when we split the shorter sample for which we have those events and are omitted for brevity.

25. The table uses a value-weighted industry return; results are the same for equal weighting.

TABLE 11 Split Sample Results: Coverage

R6	SEO	BBK	EFM	DFY1	DLTG	E/P	SIZE	INIT	OMIT
A. High Coverage									
Dependent Variable = Raw Returns									
.0772	−.0132	.0090	−.0064						
(3.83)	(2.04)	(2.04)	(.90)						
.0509	−.0152	.0088	−.0066	.0381	.0169				
(2.46)	(2.33)	(2.11)	(.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)	(.11)	(.96)		
.0518	−.0161	.0058	−.0077	.0354	.0154	.0028	.0250	.0176	.0072
(3.26)	(2.72)	(1.75)	(1.08)	(3.08)	(3.76)	(.12)	(.99)	(1.21)	(.41)
Dependent Variable = Unexpected Returns									
.0175	−.0028	.0061	−.0051						
(.81)	(.38)	(1.48)	(.69)						
.0119	−.0031	.0062	−.0052	.0053	.0027				
(.58)	(.43)	(1.53)	(.69)	(.38)	(.52)				
.0105	−.0051	.0031	−.0066	.0016	.0007	.0012	.0353		
(.63)	(.73)	(.93)	(.93)	(.15)	(.17)	(.05)	(1.31)		
.0099	−.0045	.0031	−.0065	.0021	.0007	.0014	.0360	.0139	.0076
(.60)	(.65)	(.88)	(.93)	(.19)	(.15)	(.06)	(1.35)	(.77)	(.47)

B. Low Coverage									
Dependent Variable = Raw 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)	(.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)	(.43)	(1.69)	(1.72)	(.14)
Dependent Variable = Unexpected Returns									
.0125	−.0109	.0152	−.0048						
(.70)	(.66)	(1.87)	(.26)						
.0047	−.0140	.0138	−.0079	.0237	−.0018				
(.25)	(.83)	(1.73)	(.43)	(1.62)	(.19)				
.0161	−.0014	.0115	−.0078	.0114	−.0004	.0446	−.0287		
(1.02)	(.07)	(1.51)	(.43)	(.90)	(.04)	(1.92)	(1.19)		
.0142	−.0004	.0116	−.0095	.0074	−.0001	.0426	−.0276	.0321	−.0263
(.92)	(.02)	(1.51)	(.52)	(.60)	(.45)	(1.79)	(1.11)	(1.93)	(.79)

TABLE 12 Split Sample Results: Sample Period Halves

R6	DFY1	DLTG	E/P	SIZE
A. First Half				
Dependent Variable = Raw Returns				
.0661				
(3.64)				
.0513	.0818	.0181		
(1.99)	(5.29)	(3.85)		
.0524	.0617	.0198	.0478	.0506
(2.41)	(4.47)	(4.45)	(1.77)	(2.78)
Dependent Variable = Unexpected Returns				
.0209				
(.72)				
.0057	.0283	.0038		
(.21)	(1.78)	(.78)		
.0073	.0052	.0057	.0522	.0569
(.33)	(.37)	(1.24)	(1.94)	(2.81)
B. Second Half				
Dependent Variable = Raw Returns				
.0498				
(4.42)				
.0682	.0406	.0092		
(4.18)	(2.58)	(1.83)		
.0718	.0447	.0099	.0012	(−.0290
(5.53)	(3.89)	(1.99)	(.03)	(1.39)
Dependent Variable = Unexpected Returns				
.0148				
(.90)				
.0192	− .0096	−.0022		
(1.11)	(.59)	(.38)		
.0233	−.0073	−.0016	.0022	(−.0248
(1.62)	(.64)	(.29)	(.06)	(1.20)

Note.—The table shows results from monthly Fama-MacBeth regressions for 6-month raw returns and estimated unexpected returns for the same specifications used in table 4. The top two panels use data for the independent variables from 1/1983 through 12/1990. The bottom two panels use 1/1991 through 12/1999. Standard errors are adjusted for induced 6-month autocorrelation. The resulting t statistics are in parentheses.

industry effect in our sample, amounting to over 17% for 6 months on average. (The standard errors are large because there are effectively only 20 observations in each cross section.) Interestingly, the within-industry component is also statistically very significant here, posing another test for the model.

The model passes these tests. No coefficients are significant in the bottom panel. The magnitude of the industry coefficient is 80% lower than with the raw returns and statistically not distinguishable from zero. This implies the industry effect in momentum, for all intents, is captured

TABLE 13 Industry Momentum

Industry R6	Within-Industry R6	DFY1	DLTG	E/P	SIZE
Dependent Variable = Raw Returns					
.1770	.0544				
(2.99)	(5.51)				
.0899	.0606	.0605	.0140		
(1.56)	(4.70)	(5.39)	(3.96)		
.0916	.0625	.0529	.0150	.0236	.0110
(1.78)	(5.84)	(5.90)	(4.38)	(1.05)	(.71)
Dependent Variable = Unexpected Returns					
.0372	.0184				
(.63)	(1.29)				
.0364	.0134	.0087	.0010		
(.27)	(1.00)	(.76)	(.27)		
.0406	.0157	−.0012	.0021	.0265	.0162
(.79)	(1.41)	(.14)	(.56)	(1.20)	(1.00)

Note.—The table repeats the regressions in table 4 after splitting the momentum variable R6 into industry return and within-industry return components. Industry definitions are as in Moskowitz and Grinblatt (1999) and the industry return is calculated on a value-weighted basis. Standard errors are adjusted for induced 6-month autocorrelation. The resulting t statistics are in parentheses.

entirely by industry effects in fundamental news, once this is cleanly measured.

C. Portfolio Returns

Next, we attempt to assess our model's explanatory power in terms of the returns to underreaction trading strategies. Specifically, we examine the benchmarked returns of seven anomaly-driven portfolios during our sample period. These are constructed to reflect the natural strategies based on the underreaction variables used in our regressions: buying winners, share repurchasers, or dividend initiators, and selling losers, SEO issuers, acquirers, and dividend omitters. The regression results could be misleading if the underlying relationships are nonlinear or distributional assumptions are violated.

We form portfolios each month, hold them for 6 months, and measure their cumulative abnormal performance. We apply three yardsticks for this measurement. First, we report average returns in excess of the market return to establish the degree of anomalous behavior to be explained. Second, we report returns in excess of the model's expected returns as of formation date. Third, we report each strategy's returns net of that of a random matching portfolio whose stocks have the same expected return as of the formation date. For each benchmark, we report performances for both value-weighted and equal-weighted strategies. The results are shown in table 14.

The first two columns of the table show the results adjusted only for market returns. They indicate that, apart from momentum strategies,

TABLE 14 Trading Strategy Outperformance

	Return Benchmark					
	Market Return		Expected Return		Matched Portfolio	
Strategy	EW	VW	EW	VW	EW	VW
WIN	.0044	.0026	.0006	.0000	.0006	.0003
	(.0018)	(.0020)	(.0043)	(.0042)	(.0012)	(.0019)
LOSE	−.0112	−.0060	−.0043	.0017	−.0051	.0006
	(.0027)	(.0037)	(.0056)	(.0056)	(.0019)	(.0023)
SEO	−.0016	−.0032	−.0001	−.0008	−.0005	−.0002
	(.0017)	(.0013)	(.0042)	(.0036)	(.0013)	(.0016)
BBK	.0014	.0009	.0010	.0010	.0010	.0015
	(.0013)	(.0010)	(.0035)	(.0033)	(.0007)	(.0017)
EFM	−.0003	.0004	−.0002	.0013	−.0005	.0016
	(.0017)	(.0014)	(.0038)	(.0037)	(.0012)	(.0018)
INIT	.0039	.0010	.0034	.0019	.0035	.0015
	(.0019)	(.0033)	(.0040)	(.0049)	(.0018)	(.0032)
OMIT	−.0058	.0025	−.0034	.0041	−.0038	.0041
	(.0027)	(.0028)	(.0049)	(.0043)	(.0025)	(.0026)

NOTE.—The table shows average monthly performance of seven trading strategies from 1985 through 1999. The WIN strategy purchases the top 10% of stocks on CRSP in terms of trailing 6-month return. LOSE purchases the bottom 10%. SEO, BBK, EFM, INIT, and OMIT purchase stocks having had one or more event (secondary equity offering, share buyback, equity-financed merger, and dividend initiation or omission) in the preceding 6 months. Portfolios are formed each month; all strategies hold their stocks for 6 months; and the monthly performance of a strategy is the average across the six current portfolios for that strategy. The columns EW and VW evaluate equal- and value-weighted versions of each strategy. Three benchmark methods are used. In the columns under market return, all stock returns are net of the average (which is value weighted in the VW column, and equal weighted in the EW column) of stocks in our I/B/E/S universe for each month. The expected return column nets realized returns by the ex ante value of the fitted expected return, $\alpha_0 + \alpha'_1\xi_{t-1}$, using the values estimated in Section IV and setting α_0 equal to the unconditional mean. Under matched portfolio, we match each stock to 1 of 25 portfolios formed by ranking stocks on their expected return. Each stock's return then is computed net of that realized on the matching portfolio. Standard errors are in parentheses.

underreaction anomalies look milder overall on a calendar-time basis. This is well known, as is the weakening of these effects when subjected to value weighting, which is evident here. There is still a sizable momentum anomaly, however. Returns to the SEO strategy (value weighted) and both dividend strategies (equal weighted) are statistically significant and economically meaningful at over 30 basis points per month.

The next two columns demonstrate that none of the anomalous performance survives using the model's ex ante expected return as a benchmark. Neither the top momentum decile nor the bottom decile has significant performance, nor is the difference between them significant (in fact, it has the wrong sign on a value-weighted basis). The same conclusions apply to the equal-weighted initiator and omitter portfolios (and their difference) and to the value-weighted SEO portfolio.

The unexpected performances have higher standard errors (as well as lower absolute magnitude) than the first two columns, because there is additional volatility from the market return component. This volatility is not present when we benchmark against realized returns from expected-return-matched portfolios. The two rightmost columns in the table show the resulting abnormal performances.

The net returns here are roughly the same as the unexpected returns (which implies that the model's expectations for the matching portfolios were largely accurate). The value-weighted performances remain insignificant across the board. However, the lower standard errors now imply that the loser part of the momentum strategy is significant on an equal-weighted basis (as is the difference between the dividend strategies). The asymmetry seen between winners and losers suggests a possible direction for model improvement; namely, to allow for a corresponding asymmetry in our specification of analyst behavior (the Δi equation). In fact, the degree of analyst misreaction is known to be asymmetric (Easterwood and Nutt 1999), with more underreaction to negative news. This would imply that, on average, we have underadjusted DLTG and DFY1 for negative news, leading to the pattern shown.

D. Longer Horizon Returns

This study has focused exclusively on modeling and explaining returns at a 6 month horizon. However different underreaction anomalies appear to have different characteristic time scales in terms of the persistence of return drift. Momentum patterns persist in 12-month returns but decline sharply after that. Most postevent drift studies find abnormal performance that lasts longer, in some cases as long as 3 years. Analyst revisions, on the other hand, appear to induce effects for, at most, 6 months. Can the model developed here accommodate any of these differences in persistence?

At first glance, the answer would certainly be no. The model in Section II allows for no return persistence at time scales other than 6 months. This is because we opted for a simple one-lag specification (equation [2]), which implies

$$E_t(r_{t+m}) = E_0(r_{t+m}) = \alpha_0$$

for all $m > 1$. So, current news never induces drift beyond 6 months.

We easily can generalize the model's basic structure, however. If we include further lags of our news variables on the right-hand side, then time t innovations will still enter into time $t+1$ expectations for $t+2$ returns. The most parsimonious specification assumes that the influence of lagged periods declines geometrically and may be written

$$r_t = \alpha_0 + \sum_{m=0}^{\infty} \kappa^m \alpha'_1 \xi_{t-(m+1)} + \varepsilon_t^{(0)}.$$

TABLE 15 12-Month Returns

R6	DFY1	DLTG	E/P	SIZE
Dependent Variable = Raw Returns				
.1017				
(4.92)				
.1499	.0624	.0108		
(5.53)	(2.51)	(1.62)		
.1422	.0458	.0142	.0499	.0204
(5.78)	(2.70)	(2.14)	(.95)	(.63)
Dependent Variable = Unexpected Returns				
.0601				
(1.89)				
.0769	−.0400	−.0098		
(2.67)	(1.63)	(1.36)		
.0794	−.0562	−.0071	.0484	.0192
(3.17)	(3.33)	(.95)	(.95)	(.57)

NOTE.—The table repeats the regressions in table 4 using 12-month returns as the dependent variable. The model in Section IV is extended to predict longer horizon returns via a geometric-lag specification (cf, equation [11]). Standard errors are adjusted for induced 12-month autocorrelation. The resulting t statistics are in parentheses.

This implies

$$E_t(r_{t+1+m} - \alpha_0) = \kappa^m E_t(r_{t+1} - \alpha_0),$$

so that excess return forecasts decline exponentially to their unconditional level with forecast horizon.

Fitting this generalization is straightforward as it involves only one new variable, κ, the decay rate.[26] The estimated value of this parameter is about 0.4, implying that, indeed, we should expect significant return continuation beyond 6 months.

We now repeat the basic regressions from table 4, using 12-month raw and unexpected returns as the dependent variables. The results, shown in table 15, indicate that the model has difficulty simultaneously accounting for the different decay rates of momentum returns and post-revision returns. While the 12-month unexplained momentum is insignificant on a univariate basis, it is economically large. And, when the revision variables are added, the model is able to explain momentum (to the extent it does) only at the cost of overcorrecting the revision effects.

We conclude from this that, while our informational proxy variables may summarize the shocks that cause return continuation, it does not follow that all such shocks induce the same degree of persistence. This is not surprising given prior findings in this area.

We also recomputed the returns to strategy portfolios examined previously when the holding period is extended to 12 months. Again, we

26. The solution equations change again, however. These are given in Appendix A.

TABLE 16 Trading Strategy 12-Month Outperformance

	Return Benchmark					
	Market Return		Expected Return		Matched Portfolio	
Strategy	EW	VW	EW	VW	EW	VW
WIN	.0037	.0024	.0015	.0028	.0006	.0022
	(.0017)	(.0022)	(.0043)	(.0042)	(.0011)	(.0015)
LOSE	−.0097	−.0106	−.0040	.0008	−.0038	.0011
	(.0025)	(.0038)	(.0055)	(.0054)	(.0012)	(.0022)
SEO	−.0032	−.0048	−.0022	−.0018	−.0025	−.0016
	(.0016)	(.0018)	(.0042)	(.0036)	(.0011)	(.0013)
BBK	.0019	.0002	.0017	.0020	.0015	.0023
	(.0013)	(.0012)	(.0035)	(.0034)	(.0006)	(.0016)
EFM	−.0004	−.0021	−.0005	.0012	−.0009	.0011
	(.0017)	(.0015)	(.0036)	(.0037)	(.0011)	(.0017)
INIT	.0040	.0024	.0035	.0022	.0032	.0017
	(.0019)	(.0021)	(.0040)	(.0045)	(.0015)	(.0017)
OMIT	−.0048	−.0058	−.0019	−.0001	−.0018	.0005
	(.0024)	(.0033)	(.0046)	(.0039)	(.0021)	(.0020)

Note.—The table shows average monthly performance of the seven trading strategies analyzed in table 14 when the holding period is 12 months. The benchmarks used for the columns labeled matched portfolio and expected returns use a geometric-lag versions of the model in Section IV, (cf, equation [11]) and are otherwise as described in the note to table 14. Standard errors are in parentheses.

used our geometric lag model to compute expected returns, but otherwise repeated exactly the procedure described in Section V.C. Table 16 summarizes the strategies' performance.

Compared with table 14, the average monthly performance numbers for momentum strategies go down slightly (in absolute terms) when positions are held for 6 extra months and the postevent drift numbers tend to go up. As before, all the unexpected returns are insignificant, with the exception of some of the equal-weighted strategies when benchmarked against portfolios of the same expected return. In that column (second from the right), we see evidence that (in addition to loser stocks) SEO issuers, repurchasers, and dividend initiators have significant drift at the longer horizon that the model does not explain.

This is further evidence that different events induce different degrees of persistence in returns. Our basic contention is that momentum effects stem from the aggregate influence of all such events. So, in trying to account for momentum, the model, in effect, fits an average level of persistence, which is somewhere in between the (higher) level that would best explain the postevent returns and the (lower) level that would be better for the postrevision returns.

E. Conditioning Events on Value or Glamour

For all three classes of event examined in Section IV prior research has found that postevent drift is most extreme when the company in question already appears misvalued based on its book-to-market ratio. Ikenberry

TABLE 17 Return Regressions with Events Selected by Value or Glamour

R6	gSEO	vBBK	gEFM	DFY1	DLTG	E/P	SIZE
			Dependent Variable = Raw Returns				
.0547	−.0192	.0248	−.358				
(5.11)	(1.80)	(2.60)	(2.92)				
.0585	−.0227	−.0003	−.0195	.0633	.0141		
(3.70)	(2.08)	(.02)	(1.56)	(5.02)	(3.56)		
.0592	−.0204	−.0023	−.0174	.0603	.0138	.0038	.0102
(4.61)	(2.21)	(.18)	(1.64)	(6.86)	(3.68)	(.16)	(.58)
			Dependent Variable = Unexpected Returns				
.0186	−.0072	−.0122	−.0133				
(1.07)	(.63)	(.94)	(1.00)				
.0128	−.0092	−.0125	−.0146	.0102	.0023		
(.76)	(.81)	(.95)	(1.09)	(.80)	(.56)		
.0143	−.0062	−.0148	−.0124	.0048	.0018	.0076	.0156
(1.05)	(.62)	(1.45)	(1.04)	(.55)	(.45)	(.32)	(.87)

NOTE.—The table repeats the regressions in table 8 but restricts the event sample. Now, gSEO and gEFM denote only those (“glamour”) secondary issues and mergers in which the firm falls in the bottom quintile (using NYSE breakpoints) when sorted by book-to-market ratio at the end of the month preceding the event month. Likewise, vBBK selects only (“value”) repurchase events for which the firm falls in the top quintile. The data are monthly observations from 1985 through 1999. Standard errors are adjusted for induced 6-month autocorrelation. The resulting t statistics are in parentheses.

et al. (1995) report that repurchasers in the highest book-to-market quintile average over 45% excess returns in the subsequent four years. Rau and Vermaelen (1998) sort acquiring firms in mergers by book-to-market value, and find the largest 3-year underperformance in the bottom third. Similarly, Brav, Geczy, and Gompers (2000) argue that underperformance of SEO firms is largely due to “glamour” (low B/M) issuers. Can our model explain the performance of these conditionally selected events?

We rerun the regressions reported in table 8 after selecting the subset of our events based on their book-to-market ranking in the month prior to the event. We define a “value” buyback as one in the top B/M quintile and “glamour” SEOs and acquisitions as one in the bottom quintile, using quintile breakpoints of NYSE stocks. The top panel of table 17 confirms that selecting events this way leads to more extreme drift for all three. In the first regression, for example, the coefficient on the repurchaser dummy is three times as large as the corresponding number in table 8.

However, the second panel shows that this additional drift is fully accounted for by the reaction of our news proxies. There is no unexpected drift following any of the conditional events.

The expected returns used in the regressions are computed after having purged the news proxies of analyst unconditional reaction to the event types. We could improve on this by further modeling their forecast revisions conditional on event type. It turns out analysts do not underreact at all to value buybacks (in contrast to their persistent underreaction to buybacks overall) and they do underreact to equity-financed acquisitions by glamour firms (whereas for equity-financed acquisitions generally they do not). These same patterns show up (although not significantly) in the previous regressionse: the model slightly overcorrects the returns of value repurchasers and undercorrects those of glamour acquirers. So, yet again, we see the finding, recurrent through this study, that patterns of return underreaction closely parallel those of analyst underreaction.

To summarize, there are still facets of the underreaction puzzles that the model is not able to account for. Despite its algebraic complexity, it is still oversimplified. The preceding checks give us confidence, however, that its basic conclusions are not dependent on, or sensitive to, a particular choice of sample or methodology.

VI. Concluding Remarks

This paper builds on and extends a developing body of evidence suggesting that underreaction to fundamental news is a generic phenomenon, which, in aggregate, explains momentum in stock returns. Our study complements that of Pritamani and Singal (2001), Wesley Chan (2003), and Cohen et al. (2002) in explicitly knitting together the strands of this argument. Our findings are 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.

The key step in our empirical work is the construction of good proxies for true changes in these expectations. Our approach is to cleanse analysts' forecasts of their predictable component, or at least part of it. There are a number of potentially important patterns in analyst behavior that we did not correct for, including asymmetric response to good versus bad news and more positive bias for smaller firms, firms with large intangibles, and of course, firms having investment banking relationships with the analyst's parent company. Undoing these distortions could improve our news proxies further and perhaps further extend the range of anomalies they can explain.

The word *explain* is used here in a very narrow sense, of course. We have not offered a theory of why news about fundamentals should induce return persistence. That is very much an open and important question. However we think that bringing the various underreaction anomalies into a unified framework sharpens the focus of the search for an answer.

We see room for at least two interpretations of our findings. On the one hand, it is hard not to be struck by the fact that in building our "corrected" expectational proxies from reported forecasts, we uncover the same biases in the I/B/E/S data observed in returns themselves: analysts appear to significantly underreact to 6 to 12 months of stock returns and to up to several years of corporate actions. Certainly a natural interpretation of this coincidence is that the analysts' errors, in fact, were the market's errors and 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 forecasts for well-known distortions and bringing them closer to actual expectations, then our findings suggest that returns persist because expectations do change, not because they fail to. This would imply that investors are fully aware of the impact of fundamental news and demand higher or lower returns because of it. This type of rational story faces some tough hurdles, however, in particular, identification of the exact risk for which compensation is being demanded and which is large enough to explain the magnitudes of the return differentials seen in the data.

Whatever the correct interpretation turns out to be, we believe linking return patterns to expectations (even if perhaps mistaken ones) about fundamental quantities like earnings and growth rates casts these puzzles in a simpler light than they have sometimes appeared.

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 used in Section III, which allows only one lag of forecast revisions to enter the Δi equation.

Unless otherwise specified, we assume returns are generated by equation (2),

$$r_t = \alpha_0 + \alpha_1' \boldsymbol{\xi}_{t-1} + \varepsilon_t^{(0)},$$

where $\boldsymbol{\xi}$ is the cumulative unexpected revision defined in equation (1). Let $\tilde{r}_t \equiv r_t - \alpha_0$ be the demeaned return.

Lemma A.1. Suppose analysts' revisions obey

$$\Delta i_t = \boldsymbol{\mu} + A\Delta i_{t-1} + a(L)\tilde{r}_{t-1} + \varepsilon_t^{(1)},$$

which is assumed strictly stationary. Then

$$\boldsymbol{\xi}_t = (I + EL)^{-1} F \boldsymbol{\lambda}_t,$$

where

$$F = [I - A - a(1)\alpha'_1]^{-1}$$

$$E = Fa(1)\alpha'_1$$

$$\boldsymbol{\lambda}_t = (\varepsilon_t^{(1)} + a(1)\tilde{r}_t).$$

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

$$\xi_t - A\xi_t = \varepsilon_t^{(1)} + a(1)\tilde{r}_t + a(1)\alpha'_1\xi_t - a(1)\alpha'_1\xi_{t-1} \tag{A.1}$$

or

$$\xi_t = (I - A)^{-1}\left\{ \varepsilon_t^{(1)} + a(1)[\tilde{r}_t - \alpha'_1(\xi_{t-1} - \xi_t)]\right\}. \tag{A.2}$$

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

$$\xi_t \equiv \sum_{k=0}^{\infty}[E_{\mathrm{t}}(\Delta i_{t+k}) - E_{\mathrm{t}-1}(\Delta i_{t+k})].$$

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

Now,

$$E_t(\Delta i_t) = \Delta i_t$$

$$E_t(\Delta i_{t+1}) = A\Delta i_t + a(L)\tilde{r}_t$$

$$E_t(\Delta i_{t+2}) = AE_t(\Delta i_{t+1}) + a(L)E_t(\tilde{r}_{t+1})$$

$$= AE_t(\Delta i_{t+1}) + a_0(\alpha'_1\xi_t) + a_1\tilde{r}_t + \ldots$$

$$E_t(\Delta i_{t+3}) = AE_t(\Delta i_{t+2}) + a_1(\alpha'_1\xi_t) + a_2\tilde{r}_t + \ldots$$

and so on until

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

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

$$M_t = (1 - A)^{-1}[\Delta i_t + a(1)\alpha'_1\xi_t + a(1)\tilde{r}_t + \hat{a}_1\tilde{r}_{t-1} + \hat{a}_2\tilde{r}_{t-2} \ldots + \hat{a}_p\tilde{r}_{t-p}]$$

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

$$\begin{aligned}\xi_t &= (1-A)^{-1}[\Delta i_t - \Delta i_{t-1} + a(1)\alpha'_1(\xi_t - \xi_{t-1}) \\ &\quad + a(1)r_t - a_0\tilde{r}_{t-1} - a_1\tilde{r}_{t-2}\ldots - a_p\tilde{r}_{t-p-1}] + \Delta i_{t-1} \\ &= (1-A)^{-1}[\varepsilon_t^{(1)} + a(1)\tilde{r}_t + a(1)\alpha'_1(1-L)\xi_t],\end{aligned}$$

which is the same as (13). Q.E.D.

For completeness, we note the modifications to the preceding result used in Section V.D, when returns are given by the geometric lag model

$$r_t = \alpha_0 + \sum_{m=0}^{\infty} \kappa^m \alpha'_1 \xi_{t-(m+1)} + \varepsilon_t^{(0)}$$

or

$$\alpha'_1 \xi_{t-1} = (1-\kappa L)(\tilde{r}_t - \alpha_0 - \varepsilon_t^{(0)}).$$

In this case, after expanding out each term in the expectation, M_t, the future stock return terms contribute an amount

$$a(1)\left[\left(\sum_{i=1}\alpha'_i\right)\xi_t + \left(\sum_{i=2}\alpha'_i\right)\xi_{t-1} + \left(\sum_{i=3}\alpha'_i\right)\xi_{t-2} + \ldots\right] = \frac{a(1)}{1-\kappa}(1-\kappa L)^{-1}\alpha'_1\xi_t,$$

where the equality follow from the assumed geometric form of the coefficients α_i. The difference $M_t - M_{t-1}$ then contains the expression

$$a(1)\left[\frac{\alpha'_1\xi_t}{1-\kappa} - (1-\kappa L)^{-1}\alpha'_1\xi_{t-1}\right]$$

instead of a $(1)\alpha_1{}'(\xi_t - \xi_{t-1})$, which appears in the proof of the lemma. The other terms in the expression for ξ_t are unchanged. Rearranging, this leads to the analogous forms of equations (4) to (7):

$$\xi_t = (I - \kappa L)[I + (E - \kappa I)L]^{-1}F\lambda_t,$$

where the definitions of E and λ are unchanged, but now

$$F = \left[I - A - \frac{a(1)\alpha'_1}{1-\kappa}\right]^{-1}.$$

We now return to the MA(1) model of returns (equation [2]) and treat 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 that code for the occurrence of corporate actions. These are modeled 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

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

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 that follows.

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

$$\sum_{i=1}^{p} B_i^{(1)} = 0. \tag{A.3}$$

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

$$\Delta z_t = \mu + H(L)\Delta z_{t-1} + h(L)\tilde{r}_{t-1} + \varepsilon_t. \tag{A.4}$$

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

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

Then, η consists of three stacked components, the first of which is ξ. We 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

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

subject to the restriction

$$\hat{\alpha}'_1 = [\alpha'_1\ 0\ 0]. \tag{A.5}$$

We can now state the result we are after.

Proposition A.1. Assume the system given by equation (15) is strictly stationary. Then,

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

where

$$F = [I - A^{(1)}(1) - a^{(1)}(1)\alpha'_1]^{-1}$$

$$E = Fa^{(1)}(1)\alpha'_1$$

$$\lambda_t = \left\{ \varepsilon_t^{(1)} + a^{(1)}(1)\tilde{r}_t + C^{(1)}(1)[\Delta x_t - E(\Delta x_t)] \right\}.$$

Proof. We prove this by putting the Δz system in companion form and then applying the previous lemma. 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

$$y'_t = [\Delta z'_t \ \Delta z'_{t-1} \cdots \Delta z'_{t-q}]'$$

and write equation (A.4) as

$$y_t = \bar{\mu} + Jy_{t-1} + \bar{h}(L)\tilde{r}_{t-1} + \bar{\varepsilon}_t, \tag{A.6}$$

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

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

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

so that the returns are given by

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

with $\bar{\alpha}_1$ restricted as in (A.5), $\bar{\alpha}'_1 = [\hat{\alpha}'_1 \ 0 \ 0 \ \dots \ 0]$.

Now, we can apply our lemma to the stationary AR(1) system (A.6) and deduce

$$\zeta_t = (I - J)^{-1}\left\{ \bar{\varepsilon}_t + \bar{h}(1)[\tilde{r}_t - \bar{\alpha}'_1(\zeta_{t-1} - \zeta_t)] \right\}, \tag{A.7}$$

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 need analyze only the first row of the matrix $(I-J)^{-1}$. Also the quantity in square brackets is

$$\begin{bmatrix} \{\varepsilon_t + h(1)[\tilde{r}_t - \bar{\alpha}'_1(\eta_{t-1} - \eta_t)]\} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

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 (A.7) implies

$$\eta_t = [I - H(1)]^{-1}\{\varepsilon_t + h(1)[\tilde{r}_t - \hat{\alpha}'_1(\eta_{t-1} - \eta_t)]\}. \tag{A.8}$$

We multiply this equation by $[I - H(1)]$. Since, using the restriction (A.3), $H(1)$ 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

$$\xi_t - A^{(1)}(1)\xi_t - C^{(1)}(1)\eta_t^{(3)} = \varepsilon_t^{(1)} + a^{(1)}(1)\tilde{r}_t + a^{(1)}(1)\alpha'_1\xi_t - a^{(1)}(1)\alpha'_1\xi_{t-1}. \tag{A.9}$$

Also, the third line of the same formula is simply $\eta_t^{(3)} = \varepsilon_t^{(3)}$, which is just Δx_t with the mean removed. Plugging this in to the left side of (A.9) and rearranging terms yields the desired conclusion. Q.E.D.

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.

Appendix B

Description of Data and Variable Definitions

I/B/E/S Normalizations

I/B/E/S monthly consensus forecast record fields are transformed into the two variables DFY1 and DLTG used in the text as follows.

For DFY1,

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

where t indexes the month where the forecasts are made, 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 normalizing 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,

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

where 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 and private issues of common stock by a U.S. company where the issue is not an initial public offering (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 U.S. 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 U.S. 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 the date when either the target or the acquirer makes a public announcement that it held negotiations or received a formal proposal to combine, acquire, recapitalize, or the like.

References

Abarbanell, Jeffery. 1991. Do analysts earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics* 14:147–65.
Abel, Andrew B., and Olivier Blanchard. 1986. The present value of profits and cyclical movements in investment. *Econometrica* 54:249–74.
Amemiya, Takeshi. 1985. *Advanced econometrics*. Cambridge, MA: Harvard University Press.
Asquith, Paul, and David W. Mullins. 1986. Equity issues and offering dilution. *Journal of Financial Economics* 15:61–89.
Brav, Alon, Christopher Geczy, and Paul Gompers. 2000. Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics* 56:209–49.
Chan, Louis K. C., Narasimhan Jegadeesh, and Josef Lakonishok. 1996. Momentum strategies. *Journal of Finance* 51:1681–1714.
Chan, Louis K. C., Jason Karceski, and Josef Lakonishok. 2003. The level and persistence of growth rates. *Journal of Finance* 53:643–84.
Chan, Wesley S. 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics* 70:223–60.
Cohen, Randolph, Paul Gompers, and Tuomo Vuolteenaho. 2002. Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics* 66:409–62.
Daniel, Kent, Mark Grinblatt, Sheriden Titman, and Russell Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52:1035–58.
Dechow, Patricia, Amy Hutton, and Richard Sloan. 2000. The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research* 17:1–32.
Easterwood, John C., and Stacey R. Nutt. 1999. Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *Journal of Finance* 54:1777–97.
Falk, Barry, and Bong-Soo Lee. 1998. The dynamic effects of permanent and transitory labor income on consumption. *Journal of Monetary Economics* 41:371–87.
Fama, Eugene F. 1998. Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics* 49:283–306.
Fama, Eugene F., and James D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.
Gilchrist, Simon, and Charles P. Himmelberg. 1995. Evidence on the role of cash flow for investment. *Journal of Monetary Economics* 36:541–72.
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* 55:265–95.
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.
Kadiyala, Padma, and P. Raghavendra Rau. 2004. Investor reaction to corporate event announcements: Underreaction or overreaction? *Journal of Business* 77:357–86.
LaPorta, Rafael. 1996. Expectations and the cross-section of expected returns. *Journal of Finance* 51:1715–42.
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* 52:1765–90.
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.
Moskowitz, Tobias, and Mark Grinblatt. 1999. Does industry explain momentum. *Journal of Finance* 54:1249–90.
Pritamani, Mahesh, and Vijay Singal. 2001. Return predictability following large price changes and information releases. *Journal of Banking and Finance* 25:631–56.

Quah, Danny. 1990. Permananent and transitory movements in labor income: An explanation for "excessive smoothness" in consumption. *Journal of Political Economy* 98:449–75.
Rau, P. Raghavendra, and Theo Vermaelen. 1998. Glamour, value and the postacquisition performance of acquiring firms. *Journal of Financial Economics* 49:223–53.
Spiess, Katherine D., and John Affleck-Graves. 1995. Underpreformance in long-run stock returns following seasoned equity offerings. *Journal of Financial Economics* 38:243–68.
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* 15:869–900.
Travlos, Nickolaos G. 1987. Corporate takeover bids, method of payment, and bidding firms' stock returns. *Journal of Finance* 42:943–63.