Our paper suggests a simple, recursive residuals (out-of-sample) graphical approach to evaluating the predictive power of popular equity premium and stock market time-series forecasting regressions. When applied, we find that dividend ratios should have been known to have no predictive ability even prior to the 1990s, and that any seeming ability even then was driven by only two years, 1973 and 1974. Our paper also documents changes in the time-series processes of the dividends themselves and shows that an increasing persistence of dividend-price ratio is largely responsible for the inability of dividend ratios to predict equity premia. Cochrane's (1997) accounting identity—that dividend ratios have to predict long-run dividend growth or stock returns—empirically holds only over horizons longer than 5–10 years. Over shorter horizons, dividend yields primarily forecast themselves.

1. Introduction

The use of aggregate dividend ratios to predict stock market returns or the equity premium has a long tradition in finance (Dow 1920). Dividend ratios are the total dividends paid by all stocks \( \frac{D(t)}{P(t-1)} \), divided by the total stock market capitalization, either at the beginning of the year (the dividend yield, \( \frac{P(t)}{P(t-1)} \)) or at the end of the year (the dividend-price ratio, \( \frac{P(t)}{P(t-1)} \)). The equity premium (or market premium) is the return on the stock market \( \frac{Rm(t)}{P(t-1)} \) minus the return on a short-term risk-free treasury bill \( \frac{Rf(t)}{P(t-1)} \). A typical regression specification might be

\[
[Rm(t) - Rf(t)] = \gamma_0 + \gamma_1 \cdot \left( \frac{D(t-1)}{P(t-1)} \right) + \epsilon(t). \tag{1}
\]

More recently, Ball (1978), Rozeff (1984), Shiller (1984), Campbell and Shiller (1988), and Fama and French (1988, 1989) reinvigorated this interest; (Cochrane 1997 surveys the literature). Generally, dividend ratios are found to be statistically significant predictors, especially for annual equity premia.

This paper begins by replicating these findings: After defining the variables in §2, §3 shows the well-known fact that the dividend yield predicted in-sample prior to the 1990s (even though it seems to have disappeared in the 1990s). This empirical regularity—that dividend ratios seem to predict equity returns—ranks amongst the most important findings of academic finance, and it shows no signs of subsiding (e.g., Campbell and Viceira 2002). For example, a citation search lists more than 200 published articles citing the Fama and French (1988) article alone. In turn, a number of theories have recently appeared that build on the impact of predictability. For example, Barberis (2000), Brennan et al. (1997), Campbell and Viceira (1999), Liu (1999), Lynch (2001), and Xia (2001) build models for how investors should divide their assets between stocks and bonds, based on the premise that equity premia vary in a predictable fashion.

Our paper suggests a simple graphical method to evaluate and diagnose the forecasting ability of predictive regressions. Our out-of-sample diagnostic differs from more common in-sample tests in that
forecasting regressions are themselves estimated only with then-available data: both the “conditional dividend-ratio model” (the prevailing forecasting regressions) and the “unconditional historical equity premium model” (the prevailing simple moving average) are estimated as rolling forecasts to predict one-year-ahead equity premia. Our diagnostic in Figure 3 simply graphs the difference in the respective squared prediction errors over time. Although graphing recursive residuals is not novel, the fact that it has been neglected in this literature means that some rather startling facts about predictability have been generally overlooked.

Our diagnostic shows that dividend ratios’ presumed equity premium forecasting ability was a mirage, apparent even before the 1990s. Despite good in-sample predictive ability for annual equity premia prior to 1990, §4 shows that dividend ratios had poor out-of-sample forecasting ability even then. Our diagnostic illustrates over what time periods one might imagine finding predictive ability, and makes it immediately obvious that any pre-1990 out-of-sample dividend-ratio model positive predictive ability hinged on only two years, 1973 and 1974. Thus, our paper concludes that the evidence that the equity premium has ever varied predictably with past dividend ratios has always been tenuous: A market-timing trader could not have taken advantage of dividend ratios to outperform the prevailing moving average—and should have known this. By assuming that the equity premium was “like it always has been,” a trader would have performed at least as well in most of our samples.

Our paper then delves deeper into our particular predictive variable, the dividend ratio, and why—despite good theoretical reasons—it had such poor predictive ability. Section 5 investigates a plethora of alternative specifications. Despite our best attempts, we could not detect robust out-of-sample predictive ability of the standard dividend-ratio models in any variation. This paper then investigates the reason for the discrepancy between in-sample and out-of-sample performance. It is poor parameter stability. However, if Campbell and Shiller (1988) are right, changes in the dividend processes themselves could have demanded nonstationary dividend ratios’ coefficients in explaining the equity premium. Indeed, §6 documents that dividend ratios have become more nonstationary over time, itself a phenomenon not commonly known. As of 2001, the dividend ratios themselves have practically become random walks. Consequently, we can use the Campbell and Shiller (1988) theory to instrument the dividend-ratio market premium forecasting coefficients with their own time-varying autoregression coefficient estimates. Unfortunately, despite a good theoretical justification, the instruments cannot do better than the plain dividend ratios, casting even more doubt on the theory of dividend ratios as useful stock market predictors.

This leaves us with the puzzle as to what dividend-price ratios really predict. In §7, we show that although in the early part of the sample the dividend-price ratio used to be a good predictor of dividend growth rate, in recent years the ratio’s predictive ability has shifted towards an ability to predict its own future value (higher autoregressive root of dividend-price ratio) rather than one-year-ahead equity premia or dividend growth rates. Only on horizons greater than about 5 to 10 years does the Cochrane’s (1997) accounting identity (that dividend yields have to predict long-run dividend growth or market returns) begin to dominate the self-predictive properties of the dividend yield. We believe this explains why, for most of the sample period, the predictability of stock returns over annual horizons has been weak.

Section 8 produces a data-snooped estimate of what changes in the dividend-ratio coefficients would have to look like to make the dividend ratio a useful variable. A theory predicting future equity premia with lagged dividend ratios would have to predict slowly increasing coefficients until mid-1975, followed by slowly decreasing coefficients thereafter, and finally sharply increasing coefficients post-1999. In a sense, actual estimated betas seem to show a delayed reaction to the best-fit betas. Moreover, the data-snooped coefficients often indicate that a negative coefficient is called for—not too attractive for an investor drawn to dividend-ratio models based on theoretical considerations.

Section 9 reviews other critiques of the dividend-ratio forecast. Section 10 concludes.
2. Data

Our paper relies on the well-known (value-weighted CRSP index) return on the stock market (\(Rm(t) \equiv \log((P(t) + D(t))/P(t - 1))\), where \(P\) is the stock price level; \(D\), the paid dividends; the return on three-month risk-free treasury bill (called \(rf(t)\) and obtained from Ibbotson) \(\equiv \log[1 + \text{RF}(t)]\) to compute equity premia; and on the aggregate stock market’s dividend-price ratio \(\frac{D(t)}{P(t)}\) and dividend yield \(\frac{DY(t)}{P(t)}\). We use only annual data. Dividends are computed as sum of the dividend of the last 12 months, and are not reinvested over the last year period. The data are available at mansci.pubs.informs.org/ecompanion.html\(^1\).

Table 1 provides the descriptive statistics for the series. The properties of our series are well known. The average log equity premium was 5.6% in our sample period; the average dividend yield was 4.0%.

Figure 1 plots the time series of our regressand (the equity premium) and our regressors (the dividend ratios). The latter makes it apparent that there is some nonstationarity in the dividend ratios. The dividend ratios are almost random walks, while the equity premia are almost i.i.d. Not surprisingly, the augmented Dickey and Fuller (1979) test indicates that over the entire sample period, we cannot reject that the dividend ratios contain a unit root (see Stambaugh 1999 and Yan 1999).

3. In-Sample Fit

Table 2 correlates the equity premium with the lagged dividend yield. Panel A confirms the findings in Fama and French (1988, Table 3). Prior to the 1990s, the dividend yield \(DY(t)\) had significant forecasting power, the dividend-price ratio \(DP(t)\) had acceptable forecasting power. The t-statistics, both plain and Newey-West adjusted for heteroskedasticity and autocorrelation range from 1.51 to 3.40. However, when the sample is extended into 2002, the in-sample predictive ability declines, despite inclusion of 2001 and 2002.

\(^1\) For 2002, we used hand-obtained S&P500 data, because CRSP data was not yet available. This should make little difference.
The dividend-price ratio just misses conventional statistical significance levels, while the dividend-yield ratio retains good statistical significance.²

4. Out-of-Sample Forecasts

Even a sophisticated trader could not have used the regression in Table 2 to predict the equity premium. A trader could only have used prevailing information to estimate his model, not the entire sample period. Figure 2 shows the time series of dividend-yield and dividend-price ratio coefficients when only prevailing data is used to estimate them. The figures indicate that a historical observer would have progressively lowered his assessment of the influence of the dividend yield, but progressively increased his estimate of the influence of the dividend-price ratio. (Nonstationarity of the underlying dividend model is a theme of our paper, and will be covered in more detail later.) Nevertheless, only the dividend-yield beta coefficient would have indicated to an observer a reliably nonzero coefficient for a large part of the sample period. This is due to a small indicated standard error of the estimate.

Another illustration of the changing dividend model coefficients are regression coefficients by estimation subsamples. Table 3 estimates the dividend models in different subperiods. The dividend-price ratio coefficient starts out at about zero from 1926 to 1946, increases to about 0.25 from 1946 to 1970 (with very high statistical significance), and then returns to about 0.1 post 1970 (and with no statistical sig-

² If the sample ends in 2000, both dividend ratios are insignificant.
The dividend-price ratio, ironically the weaker in-

ing data. Table 4 shows that the dividend yield failed

to outperform the unconditional mean even in the

prevailing mean (albeit not at statistically significant
levels for most of the sample period when we use
a Diebold and Mariano 1995 statistic).3 Only for the
sample period of 1946–1990, the dividend-price ratio
has a Diebold and Mariano (1995) statistic of 2.09
(p-value of 4.2%)—just statistically significant.

The main contribution of our paper to the
literature—and our suggestion for other authors predict-
ing equity premia—is our simple graphical diag-
nostic in Figure 3. It makes it easy to understand
the relative performance of the forecasting models.
Plotting the cumulative sum-squared error from the
unconditional model minus the cumulative sum-
squared error from the dividend-ratio model, a
positive value indicates that the dividend ratio has
outperformed the unconditional model so far. A posi-
tive slope indicates that the dividend ratio had lower
forecasting error than the unconditional moving aver-
age equity premium in a given year.

The figure shows that the dividend yield practi-
cally never seemed to have outperformed the uncondi-
tional forecast. The dividend-price ratio sometimes
did, but like the dividend yield, it only had two really
good predictive years prior to 1990s, 1973 and 1974.
Post-1999, the dividend yield’s negative expected rate
of return prediction finally began to outperform the
unconditional mean. Still, even with the negative
post-1999 returns, the dividend-ratio models fail to
win back their proponents.

A natural question is why Fama and French (1988),
who perform similar tests, come to different con-
clusions. The reason is their sample period, which
is indicated by the arrow in the figure. Over their
period, the slope of the line is sufficiently positive to
give the dividend ratios an edge. However, extending
the test period forward or backward yields different
conclusions.

We also computed a Diebold and Mariano (1995)
statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it

\[ \text{EP}(t) = \alpha + \beta \cdot \text{X}(t - 1) + \epsilon(t). \]

The first row of each regression is the coefficient, the second line its OLS
- statistic, and the third line its Newey-West adjusted
- statistic year by year to see when the cumulative pre-
diction error would have indicated superior dividend-
ratio performance. Naturally, for dividend yields, it
**Explanation.** These figures plot the recursive coefficient estimates (i.e., using only historically available data at each point) in a regression predicting the (log) equity premium with the (log) dividend-price ratio and (log) dividend yield, respectively. The top graph plots the DP\((t)\) coefficient, the bottom graph plots the DY\((t)\) coefficients, both obtained from univariate regressions. The bars denote plus and minus one standard deviation.

**Interpretation.** Dividend-ratio coefficients (predicting equity premia) show remarkably different patterns, depending on their numerator. The beta using dividend-price ratios has high standard errors, but low variability (scale!). It crosses zero in our sample. The beta using dividend yields is always positive, economically larger, but also continuously declining.

never could.\(^4\) The dividend-price ratios, however, ignoring the 1950s (when we had few observations), had seemingly superior sample performance in 1984, 1987, and 1990 (with DM-statistics of 2 to 2.1 (roughly equivalent to a \(t\)-statistic)). Considering that the 1973 and 1974 outliers drive this marginal significance, an observer should have at least paused. But, even not considering outliers, the same observer would not have concluded superior performance in 1985–1986, 1988–1989, and post-1990.

5. **Alternative Specifications**

We also tried numerous variations. None of these variations impact our conclusion that the out-of-sample performance has always been poor.

1. We tried reinvesting the dividends, instead of summing them. There is practically no difference.

2. We tried changes in dividend ratios, because the dividend ratio is close to stationary. These changes in dividend ratios performed worse in forecasting than the dividend ratios themselves.

3. We tried simple returns and yields, instead of log returns and yields. For \(D(t)/P(t)\) and \(D(t)/P(t-1)\)
on an out-of-sample basis the conditional prediction had an RMSE of 16.3% and 17.4%, respectively, while the unconditional prediction had an RMSE of 17.0%. Again, the unconditional model beats the dividend-yield models and performs no worse statistically than the dividend-price ratio model.

(4) We tried predicting on different horizons (monthly, quarterly, multiyearly), although annual horizons seem to have been generally agreed to have the least statistical problems and the best or close-to-best performance. Sometimes, other frequencies improve the relative performance of the unconditional model, sometimes they improve the relative performance of the dividend-yield model. Under no frequency did we find the dividend-yield model to outperform in predicting at a halfway statistically significant manner.

(5) We tried to reconcile our definitions to match exactly those of Fama and French (1988). This included using only NYSE firms, predicting stock returns (rather than premia), and a 30-year estimation window. None of these changes made any difference.

As already mentioned in our discussion of Figure 3, the only significant difference is the choice of sample period. The Fama-French out-of-sample period began just after the dividend-yield model had ended a
Figure 3  Cumulative Relative Out-of-Sample, Sum-Squared Error Performance

Explanation. This figure plots

\[
\text{Net-SSE}(T) = \sum_{t=1946}^{T} \text{SE}(t)^{\text{Prevailing mean}} - \text{SE}(t)^{\text{Dividend Model}},
\]

where \( \text{SE}(t) \) is the squared out-of-sample prediction error in year \( t \). The “uncond” \( \text{SE} \) is obtained when the prevailing up-to-date equity premium average is used to forecast the following year’s equity premium. The “dividend model” \( \text{SE} \)'s are obtained from rolling regressions with either \( \text{DY}(t-1) \) or \( \text{DP}(t-1) \) as the (sole) predictor of the following year’s equity premium. For a year in which the slope is positive (particularly, 1973, 1974, and post-1999), the dividend-ratio regression model predicted better than the unconditional average out-of-sample.

Interpretation. This figure is the key and main diagnostic proposed by our paper. Relative to the simple prevailing equity premium mean, the dividend yield shows poor predictive performance out-of-sample in the 1960s. Both the dividend yield and the dividend ratio show poor performance in the 1990s. Prior to the 1990s, both dividend ratios had only two very good years, 1973 and 1974.

10-year poor run, and ended just three years before \( \text{DP}(t) \) began deteriorating.\(^5\)

(6) We tried different “fixed number of years” estimation windows. The unconditional model typically performs better or as well as the dividend ratio models if 5 or more years are used for parameter estimation.

(7) We tried standardized forecasts to see if the regressions/means could identify years ex ante in which it was likely to perform unreliably. (In other words, we used the regression prediction standard error to normalize forecast errors.) Again, the unconditional model (its forecast also standardized by its standard deviation) beat both versions of the conditional model.

(8) We tried a convex combination of the dividend-yield model prediction and the unconditional prediction. Such a “shrunk dividend-yield model” does not produce meaningfully better forecasts than the unconditional model alone.

\(^5\) Fama and French (1988, 1989) use estimation periods of 30 years to obtain an out-of-sample estimation period from 1967 to 1986 and 1967 to 1987, respectively, which avoids some high-variance returns in the 1930s. As Fama and French point out, an investor may have recognized that the post-war period was different enough from the pre-war period to avoid using an estimated dividend regression to predict equity premia prior to 1967. Similarly, the 1990s poor out-of-sample performance occurred after the Fama and French (1988) paper was written—and we know that the in-sample relationship has recently declined.
(9) We tried forecasting with the Stambaugh (1999) correction for high serial correlation in the dividend yield. This worsens the out-of-sample performance, even though the average dividend-yield coefficient decreases by 0.05 on average in the DP(t) specification and increases by 0.006 in the DY(t) specification. For DP(t) the RMSE increases from 15.99% to 16.33%; for DY(t), the RMSE increases from 17.42% to 17.48%. Both specifications continue to perform statistically no different than the unconditional model.

(10) Earnings-price, earnings-payout ratios (Lamont 1998) or more complex measures based on analysts’ forecasts (Lee et al. 1999) similarly do not appear to predict equity premia well out of sample.

(11) We tried a similar experiment for forecasts of the equity premium using the risk-free rate. We find some in-sample predictive ability on short frequencies (one-month to one-quarter), but little in-sample predictive ability on longer frequencies (one-year). In any case, the out-of-sample predictive ability on annual horizons (RMSE of 17.75%) is considerably worse than the unconditional mean equity premium (RMSE of 16.41%). Again, we do not believe there is much predictive ability coming from the short-term interest, either.

In sum, variations on the specification and variables did not produce instances that would lead one to believe that dividend yields or other variables can predict equity premia in a meaningful way. The conditional dividend-yield DY(t) regression models predicts worse than the prevailing unconditional equity premium at least since 1946. The conditional dividend-ratio DP(t) regression models predict no better than the prevailing unconditional equity premium. The data do not support the view that dividend ratios were ever an effective forecasting tool, even though the average dividend-yield coefficient decreases by 0.05 on average in the DP(t) specification and increases by 0.006 in the DY(t) specification. For DP(t) the RMSE increases from 15.99% to 16.33%; for DY(t), the RMSE increases from 17.42% to 17.48%. Both specifications continue to perform statistically no different than the unconditional model.

6. Instrumenting the Changing Dividend-Yield Process

If the theory is correct, changes in the dividend-yield autocorrelation and in the dividend yield’s ability to predict changes in dividend growth could themselves imply changes in the dividend-yield ability to predict the equity premium. Figure 4 plots estimated regression coefficients for our three main series, using all the date up-to-date.

Annual stock market returns (Rm(t)) have had low correlation, and have recently shown outright almost no correlation with DP(t − 1). However, the other two series have changed their process parameters. The dividend growth rate (AD(t)) used to be strongly negatively correlated with DP(t − 1)—it is i.i.d. today. The dividend-price ratio (DP(t)) had only mild autocorrelation in the post-WW2 period, but it is practically a random walk today: Prices continue to be roughly a random walk with relatively high variance, while dividends have remained not only stationary but also low variance.

These process changes can be used to enhance the dividend-ratio forecasting coefficients for the equity premium. Campbell and Shiller (1988) derive the following relationship:

\[
Rm(t+1) = \log \left[ \frac{P(t+1) + D(t+1)}{P_t} \right] + \log \left[ \frac{D(t) + D(t+1)}{D(t+1)} \right] + \log \left[ \frac{D(t+1)}{D(t)} \right] + \Delta D(t+1) = \log [e^{DP(t+1)} + e^{DP(t)}] + \Delta D(t+1). \quad (2)
\]

Assume |DP(t)| follows a stationary process with mean \( \Delta D(t) = \bar{d} - d \). Expand \( f(x) = \log(1 + e^x) \) using...
Taking covariances with DP prices continue to be roughly a random walk with relatively high variance, while dividends have remained not only stationary but also low variance.

The dividend price ratio (DP) had only mild autocorrelation in post-WW2 period, but it is practically a random walk today: This is because prices continue to be roughly a random walk with relatively high variance, while dividends have remained not only stationary but also low variance.

Taylor expansion around $\bar{x}$.

$$ f[x(t+1)] \approx f(\bar{x}) + f'(\bar{x})[x(t+1) - \bar{x}] $$

$$ \log(1 + e^{DP(t+1)}) \approx \log(1 + e^{DP(t)}) $$

$$ + \left[ \frac{e^{DP(t)}}{1 + e^{DP(t)}} \right] (DP(t+1) - DP(t)). \quad (3) $$

Define $\kappa = 1/(1 + e^{DP(t)})$ and $k = -\log(\kappa) - (1 - \kappa) \log(1/\kappa - 1)$. After some algebra,

$$ Rm(t+1) \approx -\kappa \cdot DP(t+1) + DP(t) + \Delta D(t+1) + k. \quad (4) $$

Taking covariances with DP(t) and dividing by variance of DP(t),

$$ \frac{\text{Cov}(Rm(t+1), DP(t))}{\text{Var}(DP(t))} \approx 1 - \kappa \cdot \frac{\text{Cov}(DP(t+1), DP(t))}{\text{Var}(DP(t))} $$

$$ + \frac{\text{Cov}(\Delta D(t+1), DP(t))}{\text{Var}(DP(t))}. \quad (5) $$

$$ \implies \beta_{Rm(t+1), DP(t)} \approx 1 - \kappa \cdot \beta_{DP(t+1), DP(t)} + \beta_{\Delta D(t+1), DP(t)}. \quad (6) $$

Note that these approximations work only for raw returns (instead of equity premia) and dividend-price ratios (but not for dividend yields). Our new model thus uses Equation (6). Specifically, recursive forecasts are carried out for the dividend growth rate and dividend-price ratio. The betas from these regressions are then substituted into Equation (6) to obtain an instrumented beta for stock return forecast.

Figure 5 plots the time-series of naive stock return betas and the time series of instrumented CS-based betas. Note from Equation (6) that the CS return beta is driven with opposite signs by the two auxiliary betas (both of which show an upward trend from Figure 4). The CS beta has declined in recent
The recursive betas are calculated using the entire history of data available. The figure plots these betas only for the period of 1946 to 2002.

years because of a large increase in autocorrelation of dividend-price ratio. In any case, the two betas plotted in Figure 5 show that the CS betas are typically slightly lower than the ordinary betas, and more so in the 1950s and from 1975 into the early 1990s.

One can decompose changes in the predictive coefficient itself into changes in the persistence of the dividend yield and into changes in the ability of the dividend yield to predict future dividends. Differentiating Equation (6), we get

\[ \Delta \beta_{Rm(t+1), \Delta D(t+1)} = -\kappa \cdot \Delta \beta_{DP(t), \Delta D(t)} + \Delta \beta_{\Delta D(t+1), \Delta D(t)} \]

(7)

where \( \kappa \) can be calibrated to be about 0.96 (\( =1/(1+e^{-3.36}) \)) for U.S. data. That is, parameter variation in the predictive coefficient can be due to parameter variation in the dividend-yield process or in the dividend-yield vs. dividend-growth relation. Using these equations, we can run a variance decomposition of \( \Delta \beta_{Rm(t+1), \Delta D(t+1), \Delta D(t)} \) accounts for 59.3%, while \( \Delta \beta_{\Delta D(t+1), \Delta D(t)} \) accounts for 18.7% of the variation in recursive \( \Delta \beta_{Rm(t+1), \Delta D(t+1), \Delta D(t)} \) in univariate regressions (for sample 1946–2002).

Table 5 shows the results. The table shows how three models predicting the prevailing stock return (not equity premia) perform: the prevailing dividend-price ratio regression, the Campbell-Shiller instrumented dividend-price ratio regressions which explicitly take the changing process on dividend yields and dividend-growth ratio into account, and the prevail-

---

**Explanation.** This figure plots recursive beta coefficients of forecasts using dividend-price ratio as a regressor. Direct forecasts are constructed using the equation \( Rm(t+1) = \beta_0 + \beta_1 \cdot DP(t) \). Campbell-Shiller forecasts are constructed using:

\[ DP(t+1) = a_0 + a_1 \cdot DP(t) \]
\[ \Delta D(t+1) = \gamma_0 + \gamma_1 \cdot DP(t) \]
\[ \beta_1 = 1 - 0.96 \cdot a_1 + \gamma_1 \]
\[ Rm(t+1) = \beta_0 + \beta_1 \cdot DP(t) \]

The recursive betas are calculated using the entire history of data available. The figure plots these betas only for the period of 1946 to 2002.
ing unconditional stock return mean. Unfortunately, despite good theoretical justification, the forecasting ability does not improve using Campbell and Shiller (1988) identities. Having taken the theory as seriously as we could, we have thus become even more skeptical about the ability of the dividend ratios to predict equity premia.

7. The Source of Poor Predictive Ability

This leaves us with the puzzle as to what the underlying source of this poor predictive ability is. Cochrane (1997) argues that the dividend-price ratio must forecast either the future returns or the dividend growth rate. His argument relies on a modification of Equation (4). Rearranging the terms in Equation (4) and recursing forward, we obtain,

$$DP(t) = Rm(t + 1) - \Delta D(t + 1) + \kappa \cdot DP(t + 1) - k$$

$$= \sum_{i=0}^{\infty} \kappa^i [Rm(t + 1 + i) - \Delta D(t + 1 + i)] + \text{constant.}$$

The second row of this equation indeed demonstrates Cochrane’s accounting identity that dividend-price ratio must forecast either the long-run future returns or the long-run dividend growth rate. However, for finite-period-ahead prediction, the first row is more informative than the second row: The dividend-price ratio must predict either the next-period stock return, the next-period dividend growth rate, or the next period dividend-price ratio.

Table 6 estimates the finite-period Equation (8). The top panel shows that we are now almost fully capturing the components of the dividend yield over any horizon. The bottom panel decomposes the dividend yield’s predictive components: Over periods of up to 5 years, the high $R^2$ in the self-forecast column shows that the dividend yield primarily predicts itself. The Cochrane Long-Run Identity Column shows that it is only for periods of longer than 5 years that the dividend yield stops predicting itself, and instead begins predicting stock market returns and dividend growth. Figure 4 shows that the dividend-price ratio used to be a good predictor of dividend growth rate in the early part of the sample, but in recent years its predictive ability has shifted towards predicting its own future value (higher autoregressive root of dividend-price ratio) instead of either dividend growth or future returns. Again, this explains why, for most of the sample period, the predictability of stock returns with dividend-price ratios has been weak.

8. A Description of the Empirically Best Time-Varying In-Sample Coefficients

What coefficient variations produces better out-of-sample prediction? That is, what kind of dividend-yield coefficient does it take to add useful information
Table 6  Decomposition of Dividend-Yield Components Over Different Horizons

<table>
<thead>
<tr>
<th>Explain: DP(t)</th>
<th>Full identity</th>
<th>Cochrane long-run identity</th>
<th>Self-forecast</th>
<th>Cochrane identity components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizon</td>
<td>Rm(t, t + h)</td>
<td>ΔD(t, t + h)</td>
<td>DP(t + h)</td>
<td>R² %</td>
</tr>
<tr>
<td>1 year</td>
<td>0.987</td>
<td>-0.994</td>
<td>0.968</td>
<td>100.0</td>
</tr>
<tr>
<td>2 years</td>
<td>0.973</td>
<td>-0.971</td>
<td>0.937</td>
<td>99.9</td>
</tr>
<tr>
<td>5 years</td>
<td>0.927</td>
<td>-0.921</td>
<td>0.853</td>
<td>99.4</td>
</tr>
<tr>
<td>7 years</td>
<td>0.902</td>
<td>-0.904</td>
<td>0.795</td>
<td>98.7</td>
</tr>
<tr>
<td>10 years</td>
<td>0.863</td>
<td>-0.910</td>
<td>0.730</td>
<td>97.4</td>
</tr>
<tr>
<td>20 years</td>
<td>0.736</td>
<td>-0.849</td>
<td>0.545</td>
<td>89.8</td>
</tr>
<tr>
<td>Significant</td>
<td>2, 7, 10</td>
<td>1, 2, 10</td>
<td>1–10</td>
<td>20</td>
</tr>
</tbody>
</table>

Explanation. All series are described in §2 and Table 1. This table plots estimates of Equation (6) over various horizons. The dependent variable is DP(t). Being an identity, all variables in the top panel are statistically significant. In the bottom panel, estimation horizons for which the Hansen-Hodrick overlapping year t-statistics are greater than 2 are indicated in the final row for appropriate horizons. The typical standard error on the univariate DP(t + 1) ranges from 0.07 on the 1-year horizon to about 0.11 on the 20-year horizon. Univariate standard errors on Rm(t + 1) and ΔD(t + 1) are about twice that.

Interpretation. DP(t) is primarily forecasting itself over horizons up to about 5 years. It is a partial forecaster of itself over 5 to 10-year horizons, and does not forecast itself over 20-year horizons. DP(t) is primarily forecasting future market returns (and some dividend growth) over horizons greater than 10 years. It does not forecast market returns or dividend growths over horizons less than 5 years.

Given that we have failed to find any out-of-sample predictive ability of dividend yields from a sound theoretical perspective, it is useful to entertain some descriptive investigations into the time-series properties of the dividend-yield coefficient. This is the ultimate data-snooping.

Figure 6 plots the dividend-yield coefficient that perfectly fits the next out-of-sample data point, using as intercept the prevailing average equity premium up to each date. Although some of the troughs and peaks necessarily line up with the well-known stock market ups and downs, the two are not the same (due to time-series changes in the dividend yield and changes in the prevailing equity premium mean). We also overlay a five-year moving average version over the coefficient series.

The best dividend-ratio coefficient would be erratic in the pre-WW2 era, negative in the post-WW2 era, but steadily increasing until about 1975 (the oil-shock), and slowly declining post-1975. Comparing the optimal ex post beta to our ex ante beta (Figure 5), the actual betas show a significant delay relative to the ex ante.

---

6 An alternative exercise would be to subtract from the dividend yield its own prevailing average. Unfortunately, the denominator often is close to zero, which explodes the coefficients.

7 This is similar to the conclusion reached in Pesaran and Timmerman (1995).

8 Please note that the ex post beta here is about predicting equity premia, whereas the ex ante beta is about predicting stock market returns.
post best betas: The actual beta continues to drop post-WW2 and increases only post-1973, just as the optimal beta ends its increase and starts its decrease; and the actual beta drops only post-1996, long lagging an ongoing decline in the optimal beta.

There are two important remaining questions: First, why is the pattern post-WW2 and post-1975 so different? Are there regime changes (Pástor and Stambaugh 2001, Viceira 1996, Jagannathan et al. 2000); and if so, why did they lead to these particular changes in the dividend yields? Second, what should an investor do if our inference is correct that the best ex post (data-snooped) dividend yield was or is negative? Is the true relationship between dividend yields and expected returns negative, as it has been out-of-sample? Our sharp rise post-WW2 is simply back to a zero coefficient, not to a positive coefficient. Should such an investor put money into the market when the dividend ratios are low, contrary to all theory?

9. Other Dividend Ratio Critiques

Fama and French (1988) ranks among the most influential papers of the last decade, so it is not surprising that a number of other papers have pointed out concerns in using the dividend yield (or ratio) to predict equity premia and stock returns and/or introduced other variables.9 For example, Goetzmann and Jorion (1993) use a bootstrap to evaluate the in-sample predictive performance of coefficient estimates and find that the Fama and French (1988) coefficient estimates are upward biased. Nelson and Kim (1993) examine coefficient biases and come to similar conclusions. Goetzmann and Jorion (1995) find that predictability in a longer sample (since 1872) is marginal and argues that these tests are influenced by survivorship bias.

9 We must apologize to all authors whose paper we have omitted for lack of space. However, see Shanken and Lewellen (2000) for a novel interpretation of predictability arising from estimation risk.
bias. Hodrick (1992) finds that Hansen-Hodrick and Newey-West statistics are biased on a horizon that is longer than one year. Stambaugh (1999) and Yan (1999) find that near-nonstationarity in the dividend ratios biases the $t$-statistics and $R^2$. None of these entertains our simple out-of-sample naive benchmark comparison. Fama and French (1989) also use our naive benchmark, but their dividend-forecast model even seems to outperform out-of-sample relative to their in-sample performance. (It is easy to miss this evidence, because the focus in Fama and French 1989 is the addition of fixed-income variables to the dividend yield.) Independently, Lee and Swaminathan (1999) find that the dividend yield has poor out-of-sample predictive ability in competition with their value-price ratio. After inclusion of their $V/P$ measure, the dividend yield has no marginal explanatory power. Their more sophisticated model employing the $V/P$ measure can beat a “static investment allocation” model, but only mildly so. Similarly, Lee et al. (1999) find that, from 1963–1996, traditional market ratios had little (in-sample) predictive power. Ang and Bekaert (2001) derive a structural model of equity premia based on dividend yields, earnings yields, and interest rates, and find that only the last has reliable explanatory power. The closest paper to our own in pointing out poor out-of-sample power may be Bossaerts and Hillion (1999), which investigates more stringent model-selection criteria for data from a number of countries. Still, they find no out-of-sample predictability in a 6/90 to 5/95 hold-out sample, using $D(t)/P(t−1)$ as their forecaster. The closest paper to our own in pointing out the possibility of a changing market is Viceira (1996), which tests whether there is a structural break in the relation between the dividend yield and stock returns (but fails to detect one).

10. Conclusion

Our paper finds that:

(1) A figure that graphs comparative sum-squared model residuals out-of-sample (like Figure 3) can act as a powerful diagnostic for equity premium and stock price prediction.

We firmly suggest that future papers which investigate variables for their predictive market timing ability diagnose their variables using the equivalent of Figure 3 in this paper.

(2) For simple dividend-yield models predicting equity premia, our diagnostic suggests that good in-sample performance is no guarantee of out-of-sample performance. There has never been convincing evidence that dividend ratios were ever useful in predicting for investment purposes, even prior to the 1990s. Neither the dividend-yield nor the dividend-price ratio had both the in-sample and out-of-sample performance that should have lead one to believe that it could outperform the simple prevailing equity premium average in an economically or statistically significant manner. A naive market-timing trader who just assumed that the equity premium was “like it has been” would typically have outperformed a trader who employed dividend-ratio forecasting regressions.

(3) Our diagnostic further suggests that any remaining explanatory predictive ability of the dividend ratios in the post-war period prior to the 1990s was due to two years only, 1973 and 1974.

(4) Our findings are not just a matter of quibbling over proper methods to compute statistical standard errors of a test statistic. They are much more basic. When the plain dividend-yield model underperforms the unconditional mean model out-of-sample—the null hypothesis—it becomes moot. In any case, even over sample period ranges where the dividend-ratio models outpredict the mean, there is neither reliability nor economic significance.

Our paper also offered some observations as to the underlying causes of poor prediction.

(1) The primary source of poor predictive ability is parameter instability. The estimated dividend-price ratio autoregression coefficient has increased from about 0.4 in 1945 to about 0.9 in 2002.

(2) The dividend yield has failed to forecast one-year-ahead returns or dividend-growth rates, because it has primarily forecast its own change. Cochrane’s accounting identity—that dividend yields must in the long run predict either stock returns or dividend growth—only finds traction on horizons of 5–10 years or more.

(3) Instrumenting the model to account for the time-varying properties of the dividend yield and
dividend growth processes does not aid the dividend ratio in predicting stock market levels.

Acknowledgments
The authors thank Eugene Fama, Ken French, Will Goetzman, Shingo Goto, Charles Lee, Richard Roll, Pedro Santa-Clara, Eduardo Schwartz, Jay Shanken, Robert Stambaugh, and especially Luis Viceira and Tuomo O. Vuolteenaho for feedback. Finally, they are grateful especially to the editor, Ravi Jagannathan, for feedback and encouragement.

References
——, L. M. Viceira. 1999. Consumption and portfolio decisions when expected returns are time varying. Quart. J. Econom. 114(2) 433–495.

Accepted by Ravi Jagannathan; received July 24, 2001. This paper was with the authors for 1 revision.