THE JOURNAL OF FINANCE • VOL. LIV, NO. 5 • OCTOBER 1999

What is the Intrinsic Value of the Dow?

CHARLES M. C. LEE, JAMES MYERS,
and BHASKARAN SWAMINATHAN*

ABSTRACT

We model the time-series relation between price and intrinsic value as a cointegrated system, so that price and value are long-term convergent. In this framework, we compare the performance of alternative estimates of intrinsic value for the Dow 30 stocks. During 1963–1996, traditional market multiples (e.g., B/P, E/P, and D/P ratios) have little predictive power. However, a V/P ratio, where V is based on a residual income valuation model, has statistically reliable predictive power. Further analysis shows time-varying interest rates and analyst forecasts are important to the success of V. Alternative forecast horizons and risk premia are less important.

MOST FINANCIAL ECONOMISTS AGREE that a stock's intrinsic value is the present value of its expected future dividends (or cash flows) to common shareholders, based on currently available information. However, few academic studies have focused on the practical problem of measuring intrinsic value.[1] Perhaps the scant attention paid to this important topic reflects the standard academic view that a security's price is the best available estimate of intrinsic value. Consequently, many researchers regard fundamental analysis, the study of public financial information to arrive at an independent measure of intrinsic value, as a futile exercise.

The case for the equality of price and value is based on an assumption of insignificant arbitrage costs.[2] When information and trading costs are trivial, stock prices should be bid and offered to the point where they fully re-

* Lee and Swaminathan are at the Johnson Graduate School of Management, Cornell University. Myers is at the University of Washington. We appreciate the helpful comments from Bill Beaver, Peter Easton, Jim Ohlson, Stephen Penman, Jay Ritter, Terry Shevlin, Robert Shiller, René Stulz (editor), Dan Thornton, an anonymous referee, and workshop participants at the University of Chicago, Cornell University, Dartmouth College, the University of Florida, N.B.E.R. Behavioral Finance Workshop, the London Business School, the Massachusetts Institute of Technology, Northwestern University, Stanford University, the University of Utah, the University of Washington, and the 1997 Western Finance Association. Ken French kindly provided the industrial portfolio and the factor portfolio returns. Financial analyst earnings forecasts were obtained from I/B/E/S Inc. Some of this work was completed during Lee's tenure as 1995–96 Visiting Economist at the New York Stock Exchange.

[1] Exceptions we discuss later include Penman and Sougiannis (1997), Abarbanell and Bernard (1995), Frankel and Lee (1997, 1998), Kaplan and Ruback (1995), and Campbell and Shiller (1988).

[2] See Shleifer and Vishny (1997) for a discussion of the limits of arbitrage.

flect intrinsic values. However, when intrinsic values are difficult to measure and/or when trading costs are significant, the process by which price adjusts to intrinsic value requires time, and price does not always perfectly reflect intrinsic value. In such a world, a more realistic depiction of the relation between price and value is one of continuous convergence rather than static equality.[3]

Once we admit the possibility that price may diverge from value, the measurement of intrinsic value becomes paramount. Aside from an emerging set of studies in the accounting literature which we discuss later, few academic studies to date have directly addressed the many practical problems associated with implementing a comprehensive valuation model. Nor has much attention been paid to the appropriate empirical benchmark(s) for assessing alternative empirical value estimates when price itself is a noisy measure of intrinsic value.

In this study, we empirically evaluate several alternative measures for the intrinsic value of the 30 stocks in the Dow Jones Industrial Average (DJIA). As a departure from the current literature, we do not require price to equal intrinsic value at all times.[4] Instead, we model the time-series relation between price and value as a cointegrated system, so that price and value are long-term convergent.[5] In this framework, we compare alternative empirical estimates of intrinsic value using two criteria: (i) their relative ability to track price variation in the DJIA over time, and (ii) their ability to predict market returns. We show that, under reasonable assumptions, superior empirical estimates of intrinsic value can perform better on either, or even both, dimensions.

The ability to predict market returns is important in tactical asset allocation decisions, in which portfolio weights across broad asset classes are determined on the basis of their relative value. The ability to track price variations over time is also important, because it captures the responsiveness of the valuation model to economic forces that affect prices of publicly traded stocks. Valuation models with superior tracking ability can be applied with greater confidence to nontraded stocks—for example, in valuing initial public offerings, or mergers and acquisitions.

[3] Perhaps the most direct evidence on the inequality of value and price for equity securities comes from the closed-end fund literature (e.g., Lee, Shleifer, and Thaler (1991), Swaminathan (1996)). The stock prices of these funds clearly do not equal their net asset value, even though net asset values are computed and reported weekly. The evidence instead shows that the price and value of closed-end funds converge over time, so that the fund discount is mean-reverting.

[4] Not all academic studies embrace the equality of price and value. Earlier studies that question this view include Shiller (1984), Summers (1986), DeBondt and Thaler (1987), and Lakonishok, Shleifer, and Vishny (1994).

[5] Two nonstationary time series are cointegrated if they are tied together in a long-run equilibrium relation. Formally, if any linear combination of two nonstationary time series can be shown to be a stationary process, then the two time series are said to be cointegrated (Hamilton (1994), Chap. 19).

This study is related to two streams of literature in accounting and finance. First, our work extends prior studies in finance that examine the relation between market multiples such as the book-to-market ratio (B/P) or the dividend yield (D/P) and subsequent market returns (e.g., Rozeff (1984), Fama and French (1988a, 1989), Campbell and Shiller (1988), Hodrick (1992), MacBeth and Emanuel (1993), and Kothari and Shanken (1997)). These studies evaluate the predictability of market returns using simple valuation heuristics, and they focus on return forecasting rather than valuation issues.

The evidence suggests that these simple valuation heuristics may not hold in recent years. For example, the price-to-book ratio (P/B) for the Dow stocks increased from an average of approximately 1.0 in 1979 to more than 3.2 by June 1996. The dividend yield on the Dow stocks has decreased from more than 6 percent to less than 2 percent over the same time period. Whether these trends are due to structural changes (such as lower interest rates, decreased dividend payouts, or growth expectations), or are indicative of market mispricings, is difficult to determine without a more complete valuation model.

We use a richer valuation model to address this question. Our results show that time-varying interest rates are an essential part of valuation models in time-series applications. In the post-1963 period, traditional market ratios that omit time-varying interest rates (such as B/P, D/P, and E/P ratios) do not predict U.S. market returns. However, during the same time period, a V/P ratio, in which "V" is estimated using a residual income formula, has reliable predictive power for U.S. market returns. Using a vector-autoregressive (VAR) simulation technique, we show that this result is robust to the inclusion of many factors in the regression, including B/P, D/P, and E/P, as well as the short-term interest rate, the ex ante default risk premium, the ex ante term structure risk premium, and past market returns.

Our study is also related to a recent line of research in the accounting literature that explores the empirical properties of the residual-income formula. The valuation equation we implement in this paper is similar to models appearing in recent studies by Abarbanell and Bernard (1995), Frankel and Lee (1997, 1998), Penman and Sougiannis (1997), and Dechow, Hutton, and Sloan (1997). However, these empirical studies examine the ability of the model to explain cross-sectional prices and/or expected returns, whereas our investigation focuses on the time-series relation between value and price.

We provide evidence on the sensitivity of this valuation model to various key input parameters for time-series applications. Specifically, we document the effect of altering the forecast horizon (3 years to 18 years), the choice of earnings forecasting method (a historical time-series model versus a model based on analyst consensus forecasts), the choice of risk premia (a market-wide time-varying risk premium, a Fama–French one-factor industry risk premium, or a Fama–French three-factor industry risk premium), and the choice of the riskless rate (short-term T-bill yield versus the long-term Treasury bond yield).

Our results show that the inclusion of a time-varying discount rate component is essential to the success of V. Intrinsic value estimates that do not include time-varying discount rates have little predictive power for returns. The choice of the riskless rate is also important. Specifically, value estimates based on short-term T-bill rates outperform value estimates based on long-term Treasury bond rates.

Once time-varying interest rates have been incorporated into the estimation process, the addition of analyst forecasts further improves the performance of the value estimate. Specifically, we find that using I/B/E/S consensus forecasts rather than forecasts based on a time-series model enhances both the tracking ability of V and the predictive power of the V/P ratio. In contrast, the choice of alternative forecast horizons and risk premia has little effect on our results.

Finally, our findings suggest a framework for reconciling the valuation literature in accounting and the returns prediction literature in finance. Traditionally, the accounting literature has emphasized the importance of fundamental value measures that track contemporaneous returns (and prices), and the finance literature has emphasized the ability of these fundamental measures to predict future returns. We suggest that when price is a noisy proxy for intrinsic value, it is reasonable to expect better value measures to perform better on both dimensions.

The remainder of the paper is organized as follows. In Section I, we discuss the cointegration of price and value. Section II introduces the residual-income valuation model. Section III describes the data and research methodology. Section IV compares the various value proxies in terms of their ability to track the level of the Dow index over time. Section V compares the predictive ability of V/P to other market value indicators. Finally, Section VI concludes with a discussion of the implications of our analysis for the valuation literature and for the market efficiency debate.

I. Convergence of Price and Value

A stock's intrinsic value is typically defined as the present value of its expected future dividends based on all currently available information. Notationally,

$$V_t^* \equiv \sum_{i=1}^{\infty} \frac{E_t(D_{t+i})}{(1+r_e)^i}. \tag{1}$$

In this definition, V_t^* is the stock's intrinsic value at time $t, E_t(D_{t+i})$ is the expected future dividends for period $t + i$ conditional on information available at time t, and r_e is the cost of equity capital based on the information set at time t. This definition assumes a flat term-structure of discount rates.

Although V_t^* is not observable, the standard view among financial economists is that a firm's stock price (P_t) is the best available empirical proxy for V_t^*. Indeed, many studies in finance and accounting begin with the presumption that the stock price is equivalent to the present value of expected future dividends—that is, $P_t \equiv V_t^*$. Under this assumption, all changes in price represent revisions in the market's expectation about future dividends or discount rates.

In this study we consider an alternative framework in which price can deviate from intrinsic value. These deviations can occur either because of noise trading (e.g., Shiller (1984) and DeLong et al. (1990)) or because of uninformed trading in a noisy rational expectation setting (Wang (1993)).[6] The magnitude and duration of the deviations depend on the costs of arbitrage (broadly defined to include information acquisition and processing costs, as well as trading and holding costs). In the long run, arbitrage forces cause price to converge to value. However, in the short run, the costs of arbitrage may be sufficiently large to prevent this convergence from occurring instantaneously.

An implication of this framework is that P_t is merely an estimate of V_t^*, which can be compared to other empirical estimates of V_t^*. For expositional purposes, let

$$\log(P_t) = \log(V_t^*) + \epsilon_t, \tag{2a}$$

$$\log(V_t) = \log(V_t^*) + \omega_t. \tag{2b}$$

These equations express a relation between price at time t, P_t, the intrinsic value at time t, V_t^*, and an empirical estimate of intrinsic value we refer to as V_t. Specifically, the log of P_t measures the log of V_t^* with a mispricing error, ϵ_t. Similarly, V_t is an observable estimate of intrinsic value, and the log of V_t measures the log of V_t^* with a measurement error, ω_t.[7]

In this framework, the relative accuracy of alternative V measures is reflected in the time-series properties of the error term. Ideally, if V measures V^* without error, ω_t is equal to zero for all t. Short of this ideal, superior V measures are those that have ω_t terms with smaller first and second moments and faster mean-reversion. In other words, we would like to construct a V measure with as small a measurement error as possible. Specifically, we would like the error term ω_t to have mean zero, a low standard deviation, and quick mean reversion (i.e., whenever ω_t deviates from the mean, we want it to revert quickly).

[6] Price diverges from value in the noise trader context because some traders follow "pseudo-signals" (signals that have the appearance, but not the substance, of value-relevant news). Examples of pseudo-signals include the advice of Wall Street "gurus" and technical and momentum-based strategies that do not consider intrinsic values. To the extent that uninformed traders make systematic estimation errors, price can also deviate from value in a noisy rational expectation framework (Wang (1993)).

[7] We use log transformations to simplify the exposition when dealing with ratios. Note that $\log(P_t)$ and $\log(V_t)$ may each be nonstationary, but if a linear combination of these two variables is mean-reverting, then they are cointegrated.

Because ω_t is not directly observable, we must draw inferences about the relative accuracy of different V measures through the time-series properties of empirical constructs, such as V/P. Consider the difference between equations (2a) and (2b):

$$\log(V_t/P_t) = \omega_t - \epsilon_t. \tag{3}$$

This equation expresses $\log(V_t/P_t)$ as the difference between the two error terms. The time-series properties of error ϵ_t are set by market (arbitrage) forces and are not within our control. However, if P is an unbiased estimator of V^*, then ϵ_t should be mean zero. Furthermore, given arbitrage, it is reasonable to expect that ϵ_t will be mean-reverting. For instance, ϵ_t may follow an AR(1), AR(2), or a more general ARMA process. If we make the additional assumption that ϵ_t and ω_t are not perfectly correlated, then we can use the V/P ratio to evaluate alternative measures of V.[8]

This analysis suggests two dimensions along which we can evaluate alternative empirical estimates of V^*:

1. *Tracking Ability: Better intrinsic value estimates yield V/P ratios that have a lower standard deviation and a faster rate of mean-reversion.* For a given ϵ_t, a better intrinsic value estimate, V, is one that leads to a lower standard deviation for V/P. Moreover, when ω_t deviates from the mean, we want it to mean-revert quickly. Conditional on a particular correlation structure between ω_t and ϵ_t, faster mean reversion in ω_t implies faster mean-reversion in V/P.[9]
2. *Predictive Power: Better intrinsic value estimates yield V/P ratios that have greater predictive power for future returns.* In our framework, if price measures intrinsic value perfectly, in other words if $\epsilon_t = 0$ for all t, then any mean-reversion in V/P is due entirely to ω_t. Unless ω_t proxies for time-varying expected returns, V/P has no predictive power for subsequent returns. Note that if ω_t is a proxy for time-varying expected returns, even if $\epsilon_t = 0$ for all t, ω_t could still predict future returns. It is impossible to completely rule out this possibility. However, in subsequent tests, we include control variables that proxy for time-varying expected returns, including short-term interest rates, ex ante term risk, ex ante default risk, and lagged market returns.

[8] If ω_t and ϵ_t are perfectly correlated, then V_t would track P_t perfectly. Empirically, none of our value estimates fit this description.

[9] Note that faster mean-reversion in V/P is not in itself sufficient to demonstrate that V is a more accurate estimate of intrinsic value. If ω_t and ϵ_t are highly correlated, it is possible that ϵ_t and ω_t are both slowly mean-reverting, but the difference between them is quickly mean-reverting. This possibility cannot be ruled out. However, if the quick mean-reversion in V/P is due entirely to correlation between ω_t and ϵ_t, V/P will have little power to predict future returns.

Assuming ϵ_t can sometimes be nonzero, a better V estimate produces a V/P measure that is a cleaner proxy of ϵ_t. Therefore, better V estimates produce V/P ratios with greater predictive power for returns. Specifically, when price is high (low) relative to value, we would expect lower (higher) subsequent returns. In the extreme case when V measures V^* perfectly, all the mean-reversion in V/P is due to ϵ_t.

In later tests, we compare alternative empirical proxies for V^* using these two criteria.

II. The Residual-Income Valuation Model

The valuation model we use to compute a proxy for V^* is based on a discounted residual income approach sometimes referred to as the Edwards–Bell–Ohlson (EBO) valuation equation.[10] Independent derivations of this valuation model have surfaced periodically throughout the accounting, finance, and economics literature since the 1930s. Recent approaches to empirically implement the model are discussed in several papers (e.g., Bernard (1994), Abarbanell and Bernard (1995), Penman and Sougiannis (1997), Frankel and Lee (1997, 1998), and Dechow et al. (1997)). In this section, we present the basic residual income equation and briefly develop the intuition behind the model.

In a series of recent papers, Ohlson (1990, 1991, 1995) demonstrates that, as long as a firm's earnings and book value are forecast in a manner consistent with "clean surplus" accounting,[11] the intrinsic value defined in equation (1) can be rewritten as the reported book value, plus an infinite sum of discounted residual income:

$$\begin{aligned} V_t^* &= B_t + \sum_{i=1}^{\infty} \frac{E_t[NI_{t+i} - (r_e * B_{t+i-1})]}{(1+r_e)^i} \\ &= B_t + \sum_{i=1}^{\infty} \frac{E_t[(ROE_{t+i} - r_e) * B_{t+i-1}]}{(1+r_e)^i}, \end{aligned} \tag{4}$$

where B_t = book value at time t, $E_t[\cdot]$ = expectation based on information available at time t, NI_{t+i} = Net Income for period $t + i$, r_e = cost of equity capital, and ROE_{t+i} = the after-tax return on book equity for period $t + i$.

[10] The term "Edwards–Bell–Ohlson," or "EBO," is coined by Bernard (1994). Recent implementations of this formula are most often associated with the theoretical work of Ohlson (1991, 1995) and Feltham and Ohlson (1995). Earlier theoretical treatments can be found in Preinreich (1938), Edwards and Bell (1961), and Peasnell (1982). Lee (1996) discusses implementation issues and the link to Economic Value Added (EVA), as proposed by Stewart (1991).

[11] Clean surplus accounting requires that all gains and losses affecting book value are also included in earnings; that is, the change in book value from period to period is equal to earnings minus net dividends ($b_t = b_{t-1} + NI_t - DIV_t$).

Note that this equation is identical to the dividend discount model, but expresses firm value in terms of accounting numbers. It therefore relies on the same theory and is subject to the same theoretical limitations as the dividend discount model. However, practical considerations, such as the availability of earnings forecasts, makes this model easier to implement for our purposes. The model also provides a conceptual framework for thinking about the relation between accounting numbers and firm value.

Equation (4) offers several important insights for equity valuation. First, it splits firm value into two components—a measure of the capital invested (B_t) and a measure of the present value of all future residual income (the infinite sum). The term in square brackets represents the abnormal earnings (or residual income) in each future period. If a firm always earns income at a rate exactly equal to its cost of equity capital, then this term is zero, and $V_t = B_t$. However, firms whose expected ROEs are higher (lower) than r_e have firm values greater (lesser) than their book values.

This equation highlights the importance of forward-looking earnings information in equity valuation. Historical book value is an inadequate proxy for intrinsic value because it measures the historical value of invested capital, not the value of future wealth-creating activities. Historical earnings (dividends) are also an inadequate proxy for intrinsic value because they are, at best, a crude proxy for future earnings (dividends). Moreover, the value of future earnings (dividends) depends critically on the interest rate used to discount them.

Several recent studies evaluate this model's ability to explain cross-sectional prices and expected returns. Penman and Sougiannis (1997) implement variations of the model using ex post realizations of earnings to proxy for ex ante expectations. Frankel and Lee (1998) implement this model using I/B/E/S analysts' earnings forecasts. They report that the resulting V measure explains close to 70 percent of cross-sectional prices in the United States, and that the V/P ratio is a better predictor of cross-sectional returns than B/P. Frankel and Lee (1997) employ the model in an international context and find similar results in cross-border valuations.[12]

Collectively, these studies show that the residual income model can be implemented to yield intrinsic value estimates that are highly correlated with cross-sectional stock prices, both in the United States and overseas. Judging from the reported price regression R^2s, the ability of value estimates from this model to explain cross-sectional prices is comparable to the discounted cash flow results reported in Kaplan and Ruback (1995), and

[12] Two other related studies use the model in slightly different contexts. Abarbanell and Bernard (1995) use the model to address the question of market myopia with respect to short-term versus long-term earnings expectations. Botosan (1997) uses the model to derive an implicit cost of equity in her analysis of the relation between corporate disclosure and cost of capital.

much higher than those achievable using earnings, book value, or dividends alone. However, little is known about the performance of the model in tracking prices and returns over time.

III. Data and Implementation Issues

A. Model Implementation Issues

A.1. Forecast Horizons and Terminal Values

Equation (2) expresses firm value in terms of an infinite series, but for practical purposes an explicit forecast period must be specified. This limitation necessitates a "terminal value" estimate—an estimate of the value of the firm based on the residual income earned after the explicit forecasting period. We use a two-stage approach to estimate the intrinsic value: (1) forecast earnings explicitly for the next three years, and (2) forecast earnings beyond year 3 implicitly, by linearly fading the period $t + 3$ ROE to the median industry ROE by period $t + T$. By using a "fade rate," we attempt to capture the long-term erosion of abnormal ROE over time. The terminal value beyond period T is estimated by taking the period T residual income as a perpetuity. This procedure implicitly assumes no value-relevant growth in cash flows after period T.

Specifically, we compute the following finite horizon estimate for each firm:[13]

$$\hat{V}_t = B_t + \frac{(FROE_{t+1} - r_e)}{(1 + r_e)} B_t + \frac{(FROE_{t+2} - r_e)}{(1 + r_e)^2} B_{t+1} + TV, \tag{5}$$

where

B_t = book value from the most recent financial statement divided by the number of shares outstanding in the current month from I/B/E/S.

r_e = the cost of equity (discussed below).

$FROE_{t+i}$ = forecasted ROE for period $t + i$. For the first three years, this variable is computed as $FEPS_{t+i}/B_{t+i-1}$, where $FEPS_{t+i}$ is the I/B/E/S mean forecasted EPS for year $t + i$ and B_{t+i-1} is the book value per share for year $t + i - 1$. Beyond the third year, *FROE* is forecasted using a linear fade rate to the industry median ROE.

[13] The equation for $T = 3$ can be reexpressed as the sum of the discounted dividends for two years and a discounted perpetuity of period-3 earnings, thus eliminating the need for the current book value in the formulation. However, for the other two versions of the model ($T = 12$ and $T = 18$), current book value is needed to forecast ROEs beyond period $t + 3$.

$B_{t+i} = B_{t+i-1} + FEPS_{t+i} - FDPS_{t+i}$, where $FDPS_{t+i}$ is the forecasted dividend per share for year $t + i$, estimated using the current dividend payout ratio (k). Specifically, we assume $FDPS_{t+i} = FEPS_{t+i} * k$.

TV = terminal value, estimated using three different forecast horizons:

$$T = 3, TV = \frac{(FROE_{t+3} - r_e)}{(1 + r_e)^2 r_e} B_{t+2}; \tag{6}$$

$$T = 12, TV = \sum_{i=3}^{11} \frac{(FROE_{t+i} - r_e)}{(1 + r_e)^i} B_{t+i-1} + \frac{(FROE_{t+12} - r_e)}{(1 + r_e)^{11} r_e} B_{t+11}; \tag{7}$$

$$T = 18, TV = \sum_{i=3}^{17} \frac{(FROE_{t+i} - r_e)}{(1 + r_e)^i} B_{t+i-1} + \frac{(FROE_{t+18} - r_e)}{(1 + r_e)^{17} r_e} B_{t+17}. \tag{8}$$

To compute a target industry ROE, we group all stocks into the same 48 industry classifications as Fama and French (1997). The industry target ROE is the median of past ROEs from all firms in the same industry. At least five years, and up to 10 years, of past data are used to compute this median.[14]

A.2. Cost of Equity Capital

The residual income model calls for a discount rate that corresponds to the riskiness of future cash flows to shareholders. Abarbanell and Bernard (1995) and Frankel and Lee (1998) find that the choice of r_e has little effect on their cross-sectional analyses. However, our focus is on the time-series properties of the model, and for this purpose it is important to incorporate a time-varying component. We do so by computing the cost-of-equity as the sum of a time-varying riskless rate and a consistent risk premium above that riskless rate.

Collectively, the 30 DJIA firms represent approximately one-fifth of the total market capitalization of all U.S. stocks. Therefore, we use the average risk premium on the NYSE/AMEX value-weighted market portfolio as an initial proxy for the risk premium of each stock. Later, we also examine the effect of using industry-specific risk premia reestimated each month based

[14] COMPUSTAT data are not available prior to 1961. Therefore, for firm-years before 1966, we use the cost-of-equity as a proxy for the target ROE.

on a one-factor and a three-factor Fama–French (1997) model.[15] We compute each risk premium with either a short-term or a long-term risk-free rate. Depending on the choice of the risk-free rate, we generate two classes of cost-of-equity estimates:

$r_e(TB)$ = monthly annualized one-month T-bill rate + market risk premium relative to returns on the one-month T-bills $(R_m - R_{tb1})$;
$r_e(LT)$ = monthly annualized long-term Treasury bond rate + market risk premium relative to returns on long-term treasury bonds $(R_m - R_{ltb})$.[16]

For each month t beginning in April 1963, the average excess return on the NYSE/AMEX market portfolio from January 1945 to month $t - 1$ is computed, and used as an estimate of the risk premium for month t.[17] Even though we reestimate the risk premia each month, we still may not fully account for time-varying risk premia. We address this issue by adding the short-term T-bill rate, an ex ante term structure risk variable, and an ex ante default risk variable to our prediction regressions.

A.3. Explicit Earnings Forecasts

The model calls for forecasts of future earnings. In the pre-1979 period, no analyst forecasts are available, so we use a time-series model to make explicit earnings forecasts for the next three years. From January 1979 onward, we use both the time-series model and the I/B/E/S consensus forecasts. I/B/E/S analysts supply a one-year-ahead ($FEPS_{t+1}$) and a two-year-ahead ($FEPS_{t+2}$) EPS forecast, as well as an estimate of the long-term growth rate (*Ltg*). We use both $FEPS_{t+1}$ and $FEPS_{t+2}$. Additionally, we use the long-term growth rate to compute a three-year-ahead earnings forecast: $FEPS_{t+3} = FEPS_{t+2}(1 + Ltg)$.[18] These earnings forecasts, combined with the dividend payout ratio, allow us to generate explicit forecasts of future book values per share and ROEs, using clean-surplus accounting.

[15] In an earlier version of the paper, we also present results with constant risk premia of 4, 5, 6, or 7 percent. None of our key results are affected by these variations in the risk premium.

[16] The long-term treasury bond rate for 1962–1972 is constructed from CRSP bond files and for 1973–1996 it is obtained from the Lehman Brothers database. The long-term Treasury bond yields are computed as a simple average for a portfolio of Treasury bonds with approximately 20 years to maturity. The Lehman index includes all Treasury bonds with 20 or more years to maturity excluding flower bonds and foreign obligated bonds. Before 1972 there were scarcely any Treasury bond issues with 20+ year horizons. Therefore, before 1972 any Treasury bond with maturity greater than 5 years (in the CRSP bond file) is included in the long-term bond portfolio. We use only fully taxable, nonflower bonds. Callable bonds are included in the portfolio. However, because the original maturity date is no longer valid for these bonds, we use the anticipated call date as the working maturity date.

[17] Excess return is market return in excess of the one-month T-bill return, or long-term Treasury bond return.

[18] Prior to 1981, I/B/E/S does not report *Ltg*. When this variable is missing, we use the composite growth rate implicit in FY1 and FY2 to forecast FY3.

For the period before 1979, we use a time-series model to forecast $t + 1$ to $t + 3$ earnings. Specifically, we estimate the following pooled time-series cross-sectional regression for all firms in the DJIA:

$$ROE_{i,t} = \alpha + \beta ROE_{i,t-1} + \epsilon_{i,t}. \tag{9}$$

To estimate this regression, we collect annual ROE data for the Dow stocks beginning in 1945. Specifically, we estimate the regression coefficients α and β using ROE data from 1945 to two years before the calendar year containing the current month. For example, to compute V for April 1975, we fit a regression to data from 1945 to 1973. The estimated α and β coefficients from this regression are then used to forecast ROEs for the next three years. Using this technique, we generate a new intercept and slope parameter for each year.[19] From 1963 to 1996, we generate 34 sets of annual parameter estimates. The mean and standard deviation for these estimates are:

	Mean	Std. Dev.
Intercept, α	0.05	0.0052
Slope, β	0.64	0.0260

These parameters are stable over time and the average slope coefficient is close to estimates obtained by Fairfield, Sweeney, and Yohn (1996) and Dechow et al. (1997) using a larger cross section of firms.

A.4. Matching Book Value to I/B/E/S Forecasts

I/B/E/S provides monthly consensus forecasts as of the third Thursday of each month. To ensure their forecasts are current, I/B/E/S "updates" (that is, "rolls forward" by one year) the fiscal year-end of all their forecasts in the month that the actual annual earnings are announced. For example, a December year-end firm may announce its annual earnings in the second week of February. In response to the announcement, I/B/E/S forecasts for that month will be moved to the next fiscal year. This ensures that the one-year-ahead forecast is always available for the next unannounced fiscal year-end.

A particular problem arises when I/B/E/S has updated its forecast, but the company has not yet released its annual reports. Because earnings announcements precede the release of financial statements, book value of equity for the fiscal year just ended may not be available when I/B/E/S updates its forecasting year-end. To ensure that our monthly estimates are based only on publicly available information, we create a synthetic book value using the clean surplus relation. Specifically, from the month of the earnings announcement until four months after the fiscal year end, we estimate the new

[19] Evidence from other studies support using a simple AR(1) model for ROEs. Dechow et al. (1997) show that the time-series of annual ROEs is reasonably captured by an AR(1) process and that a second lag adds little predictive power. Fairfield et al. (1996) show that further decomposition of income statement items beyond lagged ROE also adds little predictive power.

book value using book value data for year $t - 1$ plus earnings minus dividends ($B_t = B_{t-1} + EPS_t - DPS_t$). From the fourth month after the fiscal year end until the following year's earnings forecast is made, we use the actual reported book value from COMPUSTAT.

A.5. Dividend Payout Ratios

To estimate the sustainable growth rate, the model calls for an estimate of the expected proportion of earnings to be paid out in dividends. We estimate this ratio by dividing actual dividends from the last fiscal year by earnings over the same time period. We exclude share repurchases due to the practical problems associated with determining the likelihood of their recurrence in future periods. For firms experiencing negative earnings, we divide the dividends paid by (0.06 * total assets) to derive an estimate of the payout ratio.[20] Payout ratios of less than zero (greater than one) are assigned a value of zero (one). We compute future book values using the dividend payout ratio and earnings forecasts as follows: $B_{t+1} = B_t + NI_{t+1}(1 - k)$, where k is the dividend payout ratio.

B. Data and Sample Description

Our sample consists of all firms that have been members of the DJIA at least once on the last day of any month between May 1963 and June 1996. Financial data on these firms are collected from the merged 1995 COMPUSTAT annual industrial file. ROE data prior to the availability of COMPUSTAT are hand-collected from Moody's Stock Guide. Stock prices and returns are collected from the 1995 Center for Research in Securities Prices (CRSP) monthly tape. For the first six months of 1996, we augment these data with information obtained from Bloomberg Investment Services.

Monthly E/P and D/P ratios are based on the COMPUSTAT earnings and dividends per share from the most recent fiscal year-end.[21] Book values per share are computed using common shareholders equity as of the most recent fiscal year-end divided by shares outstanding at the end of the month in question. For the period beginning in January 1979, consensus analysts earnings per share forecasts are obtained from I/B/E/S. During this period, I/B/E/S forecasts are available publicly as of the third Thursday of each month. We use these monthly forecasts, and the most recent financial statement data, to estimate monthly V values.

We eliminate firms with missing data items or negative book values. When a firm is eliminated, we exclude both its stock price and its value measure from the aggregate index ratio. Missing values are more common prior to 1968. After 1968, most months in our sample have 30 firms, and all have at least 24 firms.

[20] The long-run return-on-total-assets in the United States is approximately 6 percent. Hence we use 6 percent of total assets as a proxy for normal earnings levels when current earnings are negative.

[21] Using earnings and dividends from the most recent four quarters yields similar results.

We use three measures of stock returns: the monthly returns on the Dow Jones stock portfolio (DJ), the monthly returns on the S&P 500 stock portfolio, and the monthly returns on the smallest quintile of NYSE stocks based on market capitalization (SFQ1).[22] As expected, the correlation between these three return measures is high. The correlation between the S&P 500 and DJ is 0.95; the correlation between the DJ and SFQ1 is 0.81; and the correlation between the S&P 500 and SFQ1 is 0.80. For simplicity of presentation, we only report prediction results for DJ. However, results for S&P 500 and SFQ1 are similar.

Table I presents summary statistics on the stock returns and the forecasting variables. Panels A, B, and C report results for the full period (May 1963 to June 1996). The three measures of stock returns together capture a broad cross section of stock market performance. During our study period, the average monthly excess return on the Dow index is 0.42 percent (or 5.0 percent per year). Average excess returns to the S&P 500 stock index is 0.36 percent (4.3 percent per year), and the small firm index is 0.68 percent (or 8.2 percent per year). The negative autocorrelation found at long horizons (two to four years) suggests slow mean reversion in large firm stock prices (see Carmel and Young (1997) for recent evidence on this issue). In later tests, we check the robustness of our prediction regression results to the inclusion of lagged market returns.

C. Intrinsic Value Measures

We consider several measures of intrinsic value: (1) end-of-month dividend yield on the Dow Jones portfolio, DJDP, defined as the dividends from the most recent fiscal year divided by end-of-month Dow Jones portfolio value; (2) end-of-month earnings-to-price ratio on the Dow Jones portfolio, DJEP, defined as earnings from the most recent fiscal year divided by end-of-month Dow Jones portfolio value; (3) end-of-month book-to-market ratio based on the latest available book value and shares outstanding, DJBM; and (4) variations of the Dow Jones value-to-price ratio, V/P. Initially, we consider four empirical specifications of V/P, in which we vary the forecast horizon (3-year or 12-year) and discount rate (short-term T-bill or long-term T-bond).

Panels B, C, and D of Table I present descriptive statistics for our forecasting variables. Panel B shows that the autocorrelation for the traditional measures (DJDP, DJEP, and DJBM) is quite high, suggesting either nonstationarity or long-term mean-reversion. The autocorrelation for the V/P measures is somewhat lower. The use of the short-term interest rate appears to reduce the autocorrelation. Panel D presents subperiod statistics for the post-1979 time period. Recall that in the post-1979 period, V/P is computed using analyst forecasts; in the pre-1979 period, we use a time-series of historical earnings to estimate future earnings. A comparison with Panel B shows that

[22] The latter two returns are obtained from CRSP Stocks, Bonds, Bills, and Inflation (SBBI) Series. Nominal returns are converted to excess returns by subtracting the monthly Treasury bill returns, and all returns are continuously compounded.

the autocorrelation in all four V/P metrics drops in the second subperiod. Later, we show that this decline is due largely to the introduction of analyst earnings forecasts.

Figures 1 and 2 provide additional insights on the time-series behavior of these ratios. Panel A of Figure 1 depicts the dividend yield and the short-term riskless rate (one-month T-bill yield). Over this time period, there is clearly an inverse relation between these variables: when short-term rates are low (high), dividend yields are high (low). Though not unexpected, this relation highlights the need to include a time-varying interest rate component in the valuation equation. Duffee (1996) reports that, since 1983, the correlation of one-month T-bill yields with yields on other longer-term Treasury instruments have significantly declined. Accordingly, our main results are reported using both the one-month rate and a long-term rate. Using a three-month rate in place of the one-month rate yields essentially the same results.

Panel B of Figure 1 presents the P/B and P/E ratios over this time period. Like the P/D ratio, the P/B ratio has increased dramatically in the second half of the sample period. The P/E ratio rose sharply in 1992 and in the first quarter of 1993 due to the unusually large number of Dow firms reporting losses in the prior year. For example, for fiscal 1991, nine of the 30 DJIA firms reported negative earnings before extraordinary items.[23] To ensure our results are unaffected by these firms, we repeat our tests using P/E ratios only from firms with positive earnings. None of the key results or conclusions are affected.

Figure 2 presents three versions of the P/V ratio. Panel A compares the P/V ratio computed using the long-term (VP3(LT)) and short-term (VP3(TB)) interest rate. Panel B compares the P/V ratio computed using analyst forecasts of earnings (VP3(TB)) and historical time-series estimates (VHP(TB)). All three value estimates are based on the three-period ($T = 3$) model expansion. The dashed vertical line indicates January 1979, the first month when analyst forecasts became available for all firms. All the P/V ratios are more stationary and exhibit faster mean-reversion than the traditional value measures in Figure 1. Although the three P/V ratios are highly correlated, these figures show that they also experience periods of significant divergence over the sample period. Figure 2, Panel B, in particular, illustrates the additional stability introduced by analyst forecasts in the post-1978 period.

After January 1979, I/B/E/S consensus forecasts are available, so the V metric can be estimated consistently using analyst forecasts for the entire subperiod. Figure 3 presents the P/B and P/V ratios for the period from January 1979 through June 1996. The particular P/V ratio depicted is the inverse of the three-period V/P ratio computed using a short-term discount rate (VP3(TB)). Compared to P/B, P/V is more stable over time and exhibits

[23] Further investigations show the poor earnings performance for these firms is not due to accounting changes in pension liability disclosures.

Table I

Summary Statistics for Monthly Returns and Forecasting Variables

The summary statistics are computed using monthly data from May 1963 to June 1996. All returns are continuously compounded excess returns expressed in percentages. DJ, S&P 500, and SFQ1 are continuously compounded excess returns of the Dow Jones stock portfolio, S&P 500 stock portfolio, and the smallest size quintile of NYSE stocks respectively. DJDP refers to the annual dividend yield on the Dow Jones, DJEP refers to the annual earnings yield on the Dow Jones, and DJBM refers to the book-to-market ratio on the Dow Jones. VPy(TB) refers to the y period value-to-price ratio using one-month T-bill rates and VPy(LT) refers to the y period value-to-price ratio using long-term Treasury bond rates for the Dow Jones, in percentage. *Def* is the annualized end-of-month default spread in percentage. *Term* is the annualized end-of-month term spread; and TB1 is the annualized end-of-month yield on the one-month Treasury bill, all in percentage.

Panel A: Univariate Statistics for Returns—Full Period (May 1963 to June 1996)

					Sum of Autocorrelations Up to Lag					
Variable	Mean	Std. Dev.	Min	Max	1	12	24	36	48	60
DJ	0.42	4.44	−28.17	15.38	0.05	−0.08	−0.38	−0.19	−0.14	0.02
S&P 500	0.36	4.17	−24.83	14.83	0.03	−0.04	−0.45	−0.36	−0.23	−0.02
SFQ1	0.68	6.10	−35.11	23.85	0.18	0.15	−0.10	0.04	0.11	0.05

Panel B: Univariate Statistics for Forecasting Variables—Full Period (May 1963 to June 1996)

					Autocorrelation at Lag					
Variable	Mean	Std. Dev.	Min	Max	1	12	24	36	48	60
DJDP	3.88	1.02	2.15	6.75	0.97	0.71	0.57	0.51	0.38	0.21
DJEP	7.44	3.08	1.38	16.20	0.98	0.64	0.42	0.29	0.24	0.24
DJBM	66.93	21.69	31.18	124.91	0.98	0.78	0.65	0.56	0.45	0.28
VP3(TB)	70.98	18.72	36.75	138.87	0.93	0.53	0.45	0.36	0.25	0.19
VP3(LT)	59.62	12.14	36.58	105.13	0.95	0.58	0.47	0.43	0.37	0.22
VP12(TB)	73.61	19.41	38.47	144.68	0.93	0.55	0.48	0.42	0.34	0.25
VP12(LT)	62.81	13.85	38.35	109.34	0.96	0.64	0.54	0.51	0.43	0.25
Def	0.45	0.17	0.13	0.92	0.91	0.53	0.15	0.12	0.24	0.23
Term	2.32	1.46	−2.19	7.01	0.87	0.40	0.15	−0.05	0.01	0.10
TB1	6.08	2.58	2.45	16.15	0.95	0.67	0.39	0.21	0.12	0.09

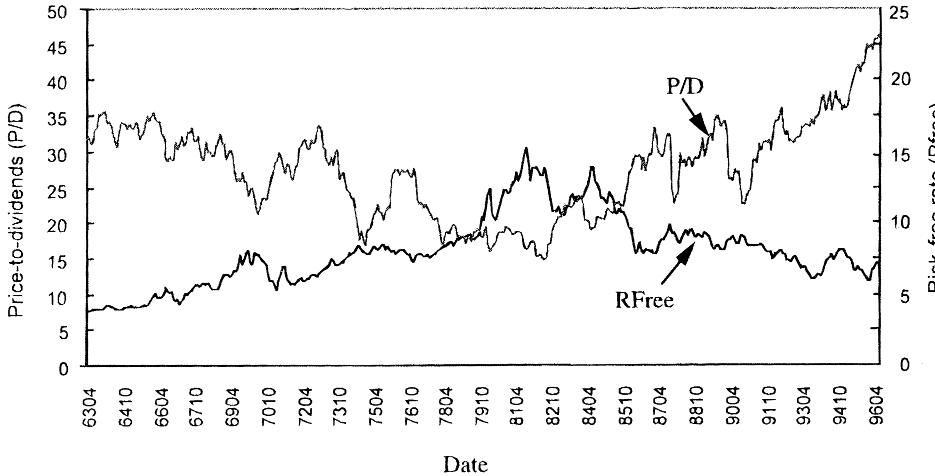
Panel C: Correlation among Forecasting Variables—Full Period (May 1963 to June 1996)

Variable	DJDP	DJEP	DJBM	VP3(TB)	VP3(LT)	VP12(TB)	VP12(LT)	*Def*	*Term*
DJEP	0.86	—	—	—	—	—	—	—	—
DJBM	0.95	0.86	—	—	—	—	—	—	—
VP3(TB)	0.65	0.51	0.64	—	—	—	—	—	—
VP3(LT)	0.75	0.73	0.77	0.89	—	—	—	—	—
VP12(TB)	0.69	0.52	0.66	0.98	0.87	—	—	—	—
VP12(LT)	0.85	0.80	0.87	0.86	0.98	0.87	—	—	—
Def	0.42	0.27	0.31	0.47	0.35	0.52	0.37	—	—
Term	0.01	−0.20	−0.02	0.60	0.21	0.58	0.17	0.37	—
TB1	0.73	0.71	0.67	0.26	0.50	0.30	0.57	0.32	−0.35

Panel D: Univariate Statistics for Forecasting Variables (January 1979 to June 1996)

					Autocorrelation at Lag					
Variable	Mean	Std. Dev.	Min	Max	1	12	24	36	48	60
DJDP	3.98	1.18	2.15	6.75	0.97	0.71	0.59	0.44	0.19	0.08
DJEP	7.49	3.58	1.38	16.20	0.97	0.62	0.35	0.13	0.01	0.07
DJBM	65.46	24.43	31.18	124.91	0.98	0.75	0.62	0.45	0.22	0.09
VP3(TB)	76.59	14.50	52.66	138.87	0.85	0.18	0.23	0.12	0.04	−0.05
VP3(LT)	62.40	9.23	42.75	90.52	0.91	0.39	0.38	0.30	0.13	0.02
VP12(TB)	80.34	16.08	52.80	144.68	0.85	0.19	0.21	0.17	0.10	0.02
VP12(LT)	65.26	11.89	45.61	100.24	0.94	0.52	0.52	0.43	0.21	0.05
Def	0.52	0.17	0.15	0.92	0.91	0.52	0.18	0.15	0.18	0.13
Term	2.76	1.53	−2.19	7.01	0.81	0.24	0.13	−0.14	−0.12	−0.07
TB1	6.93	3.04	2.45	16.15	0.94	0.68	0.43	0.23	0.10	0.04

Panel A: Price/dividend (P/D) and the riskless rate (RFree)



Panel B: Price-to-book (P/B) and price-to-earnings (P/E)

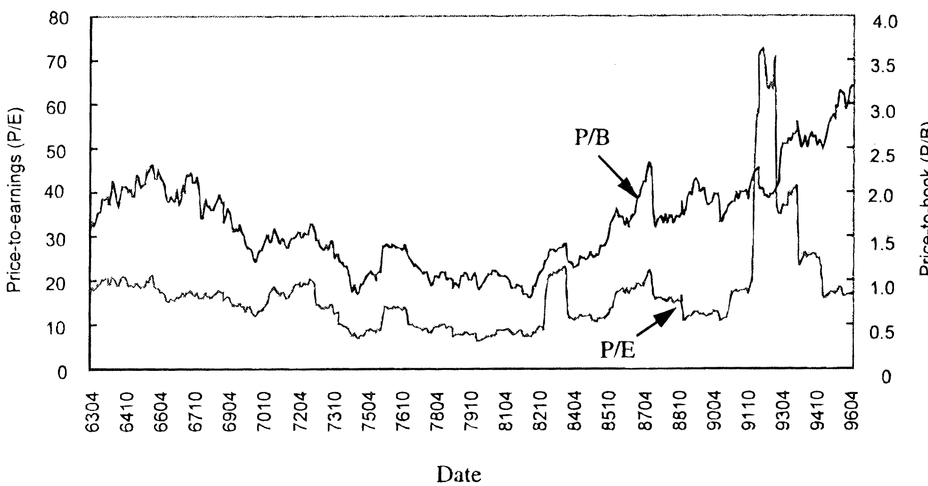


Figure 1. P/D, P/E, and P/B ratios for the DJIA (April 1963 to June 1996). These two figures depict the risk-free rate (RFree), price-to-dividends (P/D), price-to-book (P/B), and price-to-earnings (P/E) ratios for the Dow Jones Industrial Average (DJIA) stocks at monthly intervals between April 1963 and June 1996. E, B, and D represent earnings before extraordinary items, book value, and dividends respectively from the previous fiscal year end. P is the price at the end of each month. RFree is the annualized percentage yield on the 30-day T-bill as of the end of each month. Nine of the 30 DJIA stocks reported negative earnings in fiscal 1991, leading to a sharp increase in P/E during this period.

faster mean-reversion. During this time period, P/V rarely exceeds 1.8 or falls below 0.9. In fact, prior to 1996, the P/V ratio exceeds the two standard deviation mark (1.8) only twice—on November 1980, and just before the crash of September 1987 (depicted by a vertical dashed line).

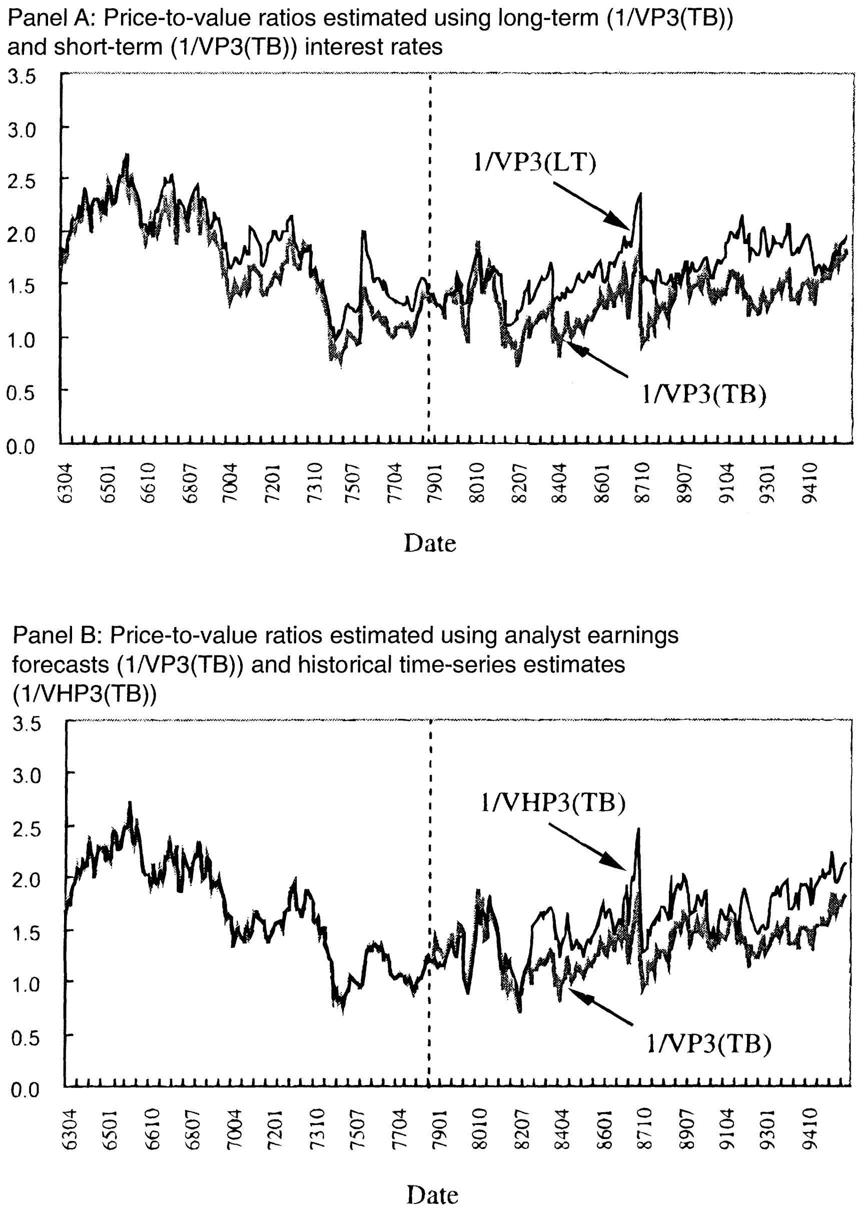


Figure 2. P/V ratios for the DJIA (April 1963 to June 1996). These two figures depict the price-to-ratio (P/V) for the Dow Jones Industrial Average (DJIA) stocks at monthly intervals between April 1963 and June 1996. P is the price at the end of each month. V is an estimate of the intrinsic value based on a residual-income model. Panel A shows P/V based on long-term (VP3(LT)) and short-term (VP3(TB)) interest rates. Panel B shows P/V based on analyst consensus earnings forecasts (VP3(TB)) and time-series estimates (VHP3(TB)). The dashed vertical line indicates the first month for which analyst forecasts are available (January 1979). The short-term riskless rate (TB) is the annualized percentage yield on the 30-day T-bill as of the end of each month. The long-term rate (LT) is the annualized percentage yield on long-term T-bonds.

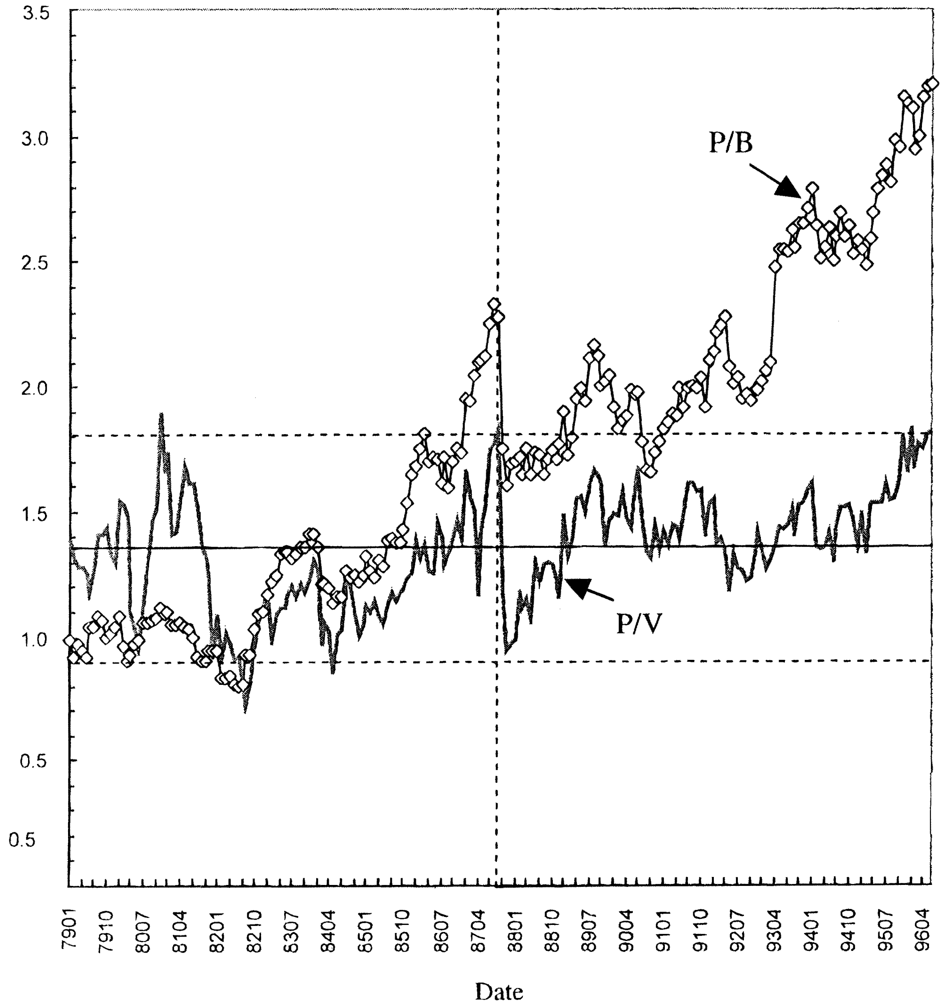


Figure 3. Price-to-book and price-to-value in recent years (January 1979 to June 1996). This graph depicts the price-to-book (P/B) and price-to-value (P/V) ratios for the 30 Dow Jones Industrial Average (DJIA) stocks at monthly intervals from January 1979 through June 1996. B represents book value from the most recent fiscal year-end divided by shares outstanding at the end of each month. V is an estimate of intrinsic value based on a three-period residual-income model using I/B/E/S analyst consensus earnings forecasts, a short-term riskless rate, and a market risk premium. Individual V and P estimates per share for each stock are aggregated to form the portfolio V and P measures. The horizontal solid line indicates the mean portfolio P/V ratio for the time period. The horizontal dotted lines indicate +/− two standard deviations based on the entire time period. The vertical line indicates September 1987.

The average P/V ratio depicted in Figure 3 is 1.35, which is significantly greater than 1. However, the average P/V is entirely scale-dependent and provides no information about market efficiency. For example, we can scale the mean P/V ratio to equal 1 by using a lower market risk premium in computing V. We do not perform this scaling because we want the discount rate used to reflect the expected market risk premia estimated from historical data. Alternatively, we can use the historical market risk premium, but extend the forecast horizon. If we select an 18-year, rather than a 3-year, forecast horizon for V, the average P/V ratio for Figure 3 would be approximately equal to 1 (see Row 6 in Table VIII). Scaling P/V in this manner makes the ratio easier to interpret, but does not improve its predictive power.

D. Business Cycle Variables

It is well known that business cycle variables such as the default spread, *Def*, and the term spread, *Term*, predict stock returns (see Fama and French (1989)). Accordingly, we control for the effects of these variables in our tests of return predictability. The default spread is a measure of the ex ante default risk premium in the economy and is measured as the difference between the end-of-month yield (annualized) on a market portfolio of corporate bonds and end-of-month yield (annualized) on a portfolio of AAA bonds. The term spread is a measure of the ex ante term risk premium in the economy and is measured as the difference between the end-of-month yield (annualized) on a portfolio of AAA bonds and the end-of-month yield on the one-month T-bill. The corporate bond yields are obtained from the Lehman Brothers corporate bond dataset and the Corporate Bond Module provided by Ibbotson Associates. The T-bill yields are obtained from CRSP bond files.

Panel C of Table I provides Pearson correlations among our forecasting variables. Note that all four V/P measures are positively correlated with the traditional value measures. The correlation is lower when V is computed using the short-term interest rate. However, using the short-term rate results in a higher correlation between V/P and the two business cycle variables (*Term* and *Def*). This suggests that the new information contained in V/P might be related to time-varying interest rates. Furthermore, V/P is positively correlated with the short-term rate (TB1). In subsequent tests, we include *Term*, *Def*, and TB1 in the predictability regressions.

IV. Tracking the Dow Index

In this section, we examine the time-series properties of our alternative intrinsic value measures. The autocorrelations in Panel B of Table I provide evidence on the time-series properties of these measures. The first-order autocorrelations of DJDP, DJEP, and DJBM are 0.97, 0.97, and 0.98, respectively. The high autocorrelations (close to one) indicate that these variable are either nonstationary or long-term mean-reverting. The half-life periods for DJDP, DJEP, and DJBM are 1.9 years, 1.9 years, and 2.9 years, respectively. This suggests that innovations to DJDP, DJEP, and DJBM take a long time to decay.

The first-order autocorrelation for the four V/P measures are smaller, suggesting that innovations to V/P lose their intensity more quickly, so that when V/P deviates from its mean, it reverts back more quickly in the subsequent months. This effect is most pronounced in the post-1978 period (Panel D), when short-term interest rates and analyst forecasts are both used to estimate V. We see that in the post-1978 period, the first-order autocorrelation for VP3(TB) and VP12(TB) is 0.85, suggesting a half-life period of approximately four months. As discussed earlier, under fairly general conditions, this evidence indicates that Vy(TB) is a better proxy for V^* than B, E, or D. Based on this benchmark, Vy(LT) is also an improvement over the traditional value metrics. However, it does not perform as well as VPy(TB).

We test the stationarity of the various intrinsic value measures more formally by conducting Phillips–Perron unit root tests (see Phillips (1987), Phillips and Perron (1988), and Perron (1988)). We run two types of Phillips–Perron unit root tests: regressions with an intercept but without a time-trend, and regressions with both an intercept and a time-trend. We do not consider the case of unit root regressions without an intercept because all the variables we are considering have nonzero means. The two types of regressions are given below:

$$\text{Without time trend:} \quad \Delta Y_t = a + (c - 1)Y_{t-1} + u_t. \tag{10}$$

$$\text{With time trend:} \quad \Delta Y_t = a + bt + (c - 1)Y_{t-1} + u_t. \tag{11}$$

The null hypothesis in both regressions is that the variable Y_t has a unit root; that is, $c = 1$. Regression (11) allows us to test the null of unit root with drift (stochastic trend) against the alternative of stationarity around a time trend. The Phillips–Perron tests allow the regression residuals to be autocorrelated and heteroskedastic. Specifically, the test procedure uses a nonparametric approach based on spectral density at zero frequency to correct for serial correlation and conditional heteroskedasticity in residuals. We report two test statistics: a regression coefficient-based test statistic, $T \times (c - 1)$, and an adjusted t-statistic based on the regression coefficient, $(c - 1)$. The adjusted t-statistics are computed allowing for serial correlation up to two lags in the regression residuals.[24]

Table II reports these two statistics based on regressions (10) and (11) for DJDP, DJEP, DJBM, VP3(TB), VP3(LT), VP12(TB), VP12(LT), *Def*, and *Term*. The results show that we cannot reject the null of unit root for DJDP, DJEP, and DJBM. This does not necessarily mean that these variables are nonstationary; however, even if they are stationary, the results show that these variables take a long time to revert to the mean. On the other hand, the null of unit root is strongly rejected for VPy(TB), *Def*, and *Term*. It is also rejected, at the 5 to 10 percent level for VPy(LT) and TB1. Again, this shows that the inclusion of time-varying interest rates produces a stationary process that mean-reverts faster than DJDP, DJEP, and DJBM.

[24] Regression results using up to 12 lags are similar and are not reported.

Table II

Phillip–Perron Unit Root Tests

This table summarizes the results of Phillips–Perron unit root tests on DJDP, DJEP, DJBM, VP3 (TB), VP3(LT), VP12(TB), VP12(LT), *Def*, *Term*, and TB1 using data from May 1963 to June 1996. DJDP refers to the annual dividend yield on the Dow Jones, DJEP refers to the annual earnings yield on the Dow Jones, and DJBM refers to the book-to-market ratio on the Dow Jones. VPy (TB) refers to the y period value-to-price ratio using 1-month T-bill rates and VPy(LT) refers to the y period value-to-price ratio using long-term Treasury bond rates for the Dow Jones, in percentage. *Def* is the annualized end-of-month default spread in percentage, *Term* is the annualized end-of-month term spread, and TB1 is the annualized end-of-month yield on the one-month Treasury bill, all in percentage. Two types of unit root tests are performed: (a) without time trend and (b) with time trend. The regression without time trend is specified as

$$\Delta Y_t = a + (c - 1)Y_{t-1} + u_t.$$

The regression with time trend is specified as

$$\Delta Y_t = a + bt + (c - 1)Y_{t-1} + u_t.$$

Two test statistics are used to test the null of unit root ($c = 1$): one is a regression coefficient based test statistic given by $T \times (c - 1)$ and the other is an adjusted t-statistic, $t(c - 1)$, corresponding to the regression coefficient $(c - 1)$. T is the number of observations. The Phillips–Perron test allows for regression errors, u_t, to be serially correlated and heteroskedastic. The test statistics are computed using serial correlation up to two lags in the regression residuals. Results using up to 12 lags are similar (not reported).

	Without Trend		With Trend		
Variable	$T * (c - 1)$	$t(c - 1)$	$T * (c - 1)$	$t(c - 1)$	T
DJDP	−7.77	−1.84	−7.98	−1.89	398
DJEP	−10.14	−2.26	−10.42	−2.31	398
DJBM	−4.61	−1.33	−5.34	−1.52	398
VP3(TB)	−25.00*	−3.57*	−28.51**	−3.75**	398
VP3(LT)	−19.77**	−3.16**	−21.06***	−3.23***	398
VP12(TB)	−24.50*	−3.53*	−28.09**	−3.69**	398
VP12(LT)	−14.54**	−2.67***	−14.65	−2.65	398
Def	−28.92*	−3.87*	−29.92*	−3.87**	398
Term	−44.61*	−4.90*	−56.02*	−5.47*	398
TB1	−18.27**	−3.12**	−18.21***	−3.08	398

*, **, *** Significant at the 1-, 5-, and 10-percent levels, respectively.

V. Returns Prediction

A. Forecasting Regression Methodology

The ability to track the value of the index does not necessarily imply an ability to predict subsequent returns. It may be, for example, that the mean reversion in V/P is due entirely to measurement errors in V. Consequently, V/P would predict changes in V but not in P. Alternatively, since both V and P measure V^* with error, it could be that these error terms are so highly positively correlated over time that V/P is not useful as a predictor of future changes in P. In this section, we evaluate the return forecasting ability of the various ratios.

The most common empirical test used in the predictability literature is the multiperiod forecasting regression test due to Fama and French (1988a, 1988b, 1989). In this regression, the average return over the next K periods is regressed on one or more explanatory variables from the current period. Consider the following OLS regression:

$$\sum_{k=1}^{K} r_{t+k}/K = X_t\theta + u_{t+K,t}. \tag{12}$$

where r_{t+k} is the continuously compounded real return per month defined as the difference between the monthly continuously compounded nominal return and the monthly continuously compounded inflation rate. X_t is a $1 \times m$ row vector of explanatory variables (including the intercept), θ is an $m \times 1$ vector of slope coefficients, K is the forecasting horizon, and $u_{t+K,t}$ is the regression residual.

The multiperiod forecasting regression may be run using either overlapping observations or nonoverlapping observations. Campbell (1993) shows that using overlapping observations increases the power of the regression to reject the null of no predictability.[25] Therefore, it is conventional to use overlapping observations for $K > 1$. However, the use of overlapping observations induces serial correlation in the regression residuals. Specifically, the regression residuals are autocorrelated up to lag $K - 1$ under both the null hypothesis of no predictability and alternate hypotheses that fully account for time-varying expected returns.[26] The regression standard errors are too low if they are not corrected for this induced autocorrelation. Moreover, the regression residuals are likely to be conditionally heteroskedastic. We correct for both the induced autocorrelation and the conditional heteroskedasticity using the Generalized Method of Moments (GMM) standard errors with the Newey–West correction (see Hansen (1982), Hansen and Hodrick (1980), and Newey and West (1987)).

We repeat these regressions for different horizons: $K = 1$, 3, 6, 9, 12, and 18. However, because the forecasting regressions at various horizons use the same data, the regression slopes will be correlated. Therefore, it is inappropriate to derive conclusions about overall predictability from the significance of any individual regression. To handle this problem, we compute the average slope statistic—that is, the arithmetic average of regression slopes at different horizons—as suggested by Richardson and Stock (1989), to test

[25] The increased power comes from two sources: (i) the average return over a longer horizon provides a better proxy of conditional expected returns than the average return over a shorter horizon, and (ii) the regression standard errors at longer horizons tend to get smaller due to the negative correlation between future expected returns and current unexpected stock returns (see Campbell (1993) for more discussion).

[26] The regression residual may be autocorrelated beyond lag $K - 1$ under the alternative hypothesis if the explanatory variables do not fully account for all of the time variation in expected returns.

the joint null hypothesis that the slopes at various horizons are equal to zero. The statistical significance of the average slope estimate is inferred from a Monte Carlo simulation, which is described below.

We run four sets of forecasting regressions: (1) univariate regressions of stock returns on each of the intrinsic value measures; (2) multivariate regressions of stock returns on all the intrinsic value measures; (3) multivariate regressions of stock returns on V/P, TB1, *Def*, and *Term*; and (4) multivariate regressions of stock returns on V/P and lagged returns. The first test evaluates the predictive power of each intrinsic value measure on its own. The second test examines the predictive power of V/P in the presence of DJDP, DJEP, and DJBM. The third test examines the predictive power of V/P in the presence of the business cycle variables, *Def*, *Term*, and TB1. Finally, we examine the predictive power of V/P controlling for lagged stock returns.

As noted earlier, asymptotic Z-statistics are computed using the GMM standard errors. However, these Z-statistics, though consistent, are likely to be biased in small samples. The bias arises from three sources. First, the independent variables in the OLS regressions are predetermined but not strictly exogenous because the regressors are a function of current price. As a result, the OLS estimators of the slope coefficients are biased in small samples (see Stambaugh (1986)). Second, as observed by Richardson and Smith (1991), although the GMM standard errors consistently estimate the asymptotic variance-covariance matrix, these standard errors are biased in small samples due to the sampling variation in estimating the autocovariances.

Third, as noted by Richardson and Stock (1989), the asymptotic distribution of the OLS estimators may not be well behaved if the degree of overlap is high relative to the sample size—that is, if K is large relative to T. Furthermore, the Z-statistics may be further biased in small samples. As a result, the null hypothesis of no predictability tends to be rejected too often. To avoid these problems, we generate small sample distributions of the OLS regression statistics using a Monte Carlo simulation (see Hodrick (1992) and Swaminathan (1996)). The Appendix describes our Monte Carlo simulation methodology.[27]

B. Forecasting Regression Results

In this section we discuss the results of our forecasting regressions. First, we discuss univariate regression results. Next, we discuss multivariate regression results involving all of the intrinsic value measures. Third, we discuss multivariate regression results involving V/P and the business cycle variables. Finally, we discuss multivariate regression results involving V/P and lagged returns.

[27] The VAR simulation methodology we employ is a more general version of the VAR found in Kothari and Shanken (1997) and Nelson and Kim (1993). Specifically, our procedure takes into account the contemporaneous correlation among the various explanatory variables and imposes the null of no predictability only on stock returns.

B.1. Univariate Regressions

The specification for the univariate forecasting regression is as follows:

$$\sum_{k=1}^{K} Y_{t+k}/K = a + bX_t + u_{t+K,t}. \tag{13}$$

We run univariate regressions for Y = DJ and X = DJDP, DJEP, DJBM, and VP$y(x)$, where y = 3 or 12 and x = TB or LT. If the stock price is too low relative to intrinsic value, then current ratios of intrinsic value to stock price (DJDP, DJDP, DJBM, and V/P) will be high. At the same time, because price reverts to value, we expect future stock returns to be high. Therefore, we expect high DJDP, DJEP, DJBM, or V/P to predict high stock returns. Thus, for all regressions, a one-sided test of the null hypothesis is appropriate.

Table III presents univariate regression results for predicting returns of the Dow 30 stocks. The column labeled *Bias* represents the mean of the empirical distribution of b generated under the null hypothesis of no predictability from 5,000 trials of a Monte Carlo simulation. The positive bias reflects a tendency for b to be positive even when the dependent variable has no ability to predict returns. $Z(b)$ is the asymptotic Z-statistic corrected for both induced autocorrelation and conditional heteroskedasticity using GMM standard errors with the Newey–West (1987) correction. To test whether these Z-statistics are significant in small samples, Table III also presents fractiles of the empirical distribution of $Z(b)$. The regressions are run for horizons K = 1, 3, 6, 9, 12, and 18 months. Note that the fractiles of statistics increase as a function of K, suggesting that the small sample inference problem is more severe at longer holding periods.

The evidence in Panel A of Table III shows that the VP3(LT) ratio has significant predictive power for Dow returns. The Z-statistics for VP3(LT) are significant at the 5 percent level at all horizons. The average slope statistic is significant at the 1 percent level. The R^2s from the regressions (from 1.6 percent to 13.6 percent) indicate that VP3(LT) is able to explain a substantial portion of future DJ returns. The slope coefficients are uniformly positive, indicating that high V/P predicts high stock returns. The average estimated slope coefficient is 0.04, indicating that a 1 percent increase in V/P results in a four basis point increase in average expected returns over the next nine months.[28]

Panel B shows that replacing the long-term rate with the short-term rate strengthens the predictive power of V/P. Though the average slope coefficient actually decreases slightly (from 0.40 to 0.33), the Z-statistics for VP3(TB) are generally higher and are significant at the 1 percent level for all hori-

[28] To ensure that the predictive power of V/P does not derive solely from the October 1987 crash, we first reestimate the prediction regression with an indicator variable for October 1987. This procedure increases the R^2 and has virtually no effect on the estimated slope coefficients and Z-statistics for V/P. We also reestimate the regression omitting the five months immediately around the crash (August to December 1987). The prediction results are similar.

zons. The regression R^2s are also higher, ranging from 3.1 percent to 20.5 percent. Panels C and D show these results carry over to the 12-period model. Increasing the forecast horizon from 3 to 12 periods had little effect on the predictive power of V/P.

The valuation literature generally recommends the use of longer term interest rates. In theory, long-term (versus short-term) rates should better reflect the opportunity cost of capital over the forecasting horizon for future cash flows. Panel A shows that incorporating long-term interest rates clearly improves returns prediction. However, Panel B shows that short-term rates perform even better. This result implies that we may be able to derive better empirical estimates of intrinsic value by using short-term rates. However, given our limited sample, we recommend caution in generalizing this result to other firms and time periods.

In contrast, the last three rows of Table III show that DJDP, DJEP, and DJBM have little predictive power for the Dow returns. Even though the average slope coefficients for all three variables have the right sign, the Z-statistics are small and the R^2s are low. Comparing the average estimated slope coefficient b for these variables to the fractiles generated from the Monte Carlo simulation, it is clear that DJDP, DJEP, and DJBM have no reliable predictive power for returns over our sample period.

Although V/P is estimated using only the Dow 30 stocks, we find its predictive power is not limited to Dow returns. In fact, we find similar or stronger results for the prediction of excess returns on the S&P 500, and on a portfolio of small firms (not reported). Specifically, we find that V/P has significant predictive power for excess returns to the S&P 500 portfolio, but that DJDP, DJEP, and DJBM do not. Similarly, we find that V/P also has strong predictive power for excess returns on a portfolio consisting of firms in the smallest NYSE size quintile.

Our results are not inconsistent with prior studies. MacBeth and Emanuel (1993) find that much of the apparent predictability of DJBM and DJDP disappear when the statistical biases discussed above are accounted for. Using more extensive annual data, Kothari and Shanken (1997) show that DJBM and DJDP have some predictive power for overall market returns in the 1926–1991 period. However, they find that the predictive power of these variables is much weaker in the 1941–1991 subperiod. Our findings add to these prior studies, and suggest that in the most recent 34 years (1963 to 1996), traditional market ratios that exclude time-varying interest rates have little or no power to predict market returns.

B.2. Multivariate Regressions Involving DJDP, DJEP, DJBM, and V/P

In this section, we report multivariate forecasting regression results involving all four measures of intrinsic value. Specifically, we run multivariate regressions of the following form:

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + b\,DJDP_t + c\,DJEP_t + d\,DJBM_t + e\,X_t + u_{t+K,t}, \qquad (14)$$

Table III

Univariate Forecasting Regressions

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + bX_t + u_{t+k,t}$$

This table summarizes univariate forecasting regression results using monthly data from May 1963 to June 1996. The dependent variable in these regressions (DJ) is the excess return on the DJIA portfolio. The independent variable (X) changes in each panel. VPy(LT) refers to the y period value-to-price ratio using long-term T-bond rates, and VPy(TB) refers to the y period value-to-price ratio using 1-month T-bill rates, in percentage. In forecasting beyond one-month ($K > 1$), the regressions use overlapping observations. b is the slope coefficient from the OLS regressions. $Z(b)$ is the asymptotic Z-statistic computed using generalized method of moments (GMM) standard errors with Newey–West correction. These standard errors correct for induced autocorrelation in regression residuals due to overlapping observations and for generalized conditional heteroskedasticity. Adj. R^2 refers to the adjusted coefficient of determination from the OLS regression. Bias refers to the mean of the empirical distribution of b generated under the null hypothesis of no predictability from 5,000 trials of a Monte Carlo Simulation. Avg. b is the average slope statistic. The columns labeled Fractiles represent the empirical distribution of $Z(b)$ and Avg. b obtained under the null hypothesis from the same Monte Carlo simulation. The last three rows, titled Average, report the average slope and fractiles of the average slope.

K	b	Bias	$Z(b)$	0.01	0.05	0.10	0.90	0.95	0.99	Adj. R^2	N
					Panel A: X = VP3(LT)						
					Fractiles of Z-Statistics						
1	0.050	0.002	2.303**	−2.243	−1.552	−1.192	1.386	1.755	2.435	1.62	398
3	0.045	0.002	2.575**	−2.706	−1.881	−1.389	1.687	2.136	2.896	4.09	396
6	0.042	0.002	2.907**	−2.824	−1.939	−1.457	1.756	2.274	3.098	7.31	393
9	0.039	0.002	3.212**	−2.909	−2.012	−1.491	1.840	2.342	3.261	9.89	390
12	0.037	0.002	3.285**	−2.959	−2.038	−1.521	1.880	2.411	3.409	12.58	387
18	0.030	0.002	3.033**	−3.229	−2.180	−1.592	2.022	2.557	3.638	13.58	381
					Fractiles of Average b						
Avg. b	0.040*			−0.029	−0.020	−0.014	0.019	0.025	0.036		

Panel B: X = VP3(TB)											
				Fractiles of Z-Statistics							
1	0.043	0.002	3.464*	−2.218	−1.514	−1.162	1.423	1.789	2.423	3.10	398
3	0.038	0.002	3.882*	−2.717	−1.859	−1.390	1.754	2.183	3.035	7.05	396
6	0.032	0.002	3.696*	−2.837	−1.884	−1.431	1.873	2.272	3.154	10.47	393
9	0.031	0.002	3.932*	−2.932	−1.940	−1.439	1.894	2.383	3.305	15.42	390
12	0.030	0.002	3.940*	−2.996	−1.994	−1.473	1.940	2.486	3.482	19.34	387
18	0.024	0.002	3.657**	−3.140	−2.086	−1.533	2.049	2.622	3.803	20.48	381
				Fractiles of Average b							
Avg. b	0.033*			−0.020	−0.014	−0.010	0.015	0.019	0.026		
Panel C: X = VP12(LT)											
				Fractiles of Z-Statistics							
1	0.034	0.003	1.839**	−2.229	−1.572	−1.230	1.428	1.772	2.439	0.89	398
3	0.031	0.003	2.021***	−2.769	−1.911	−1.460	1.730	2.217	2.970	2.45	396
6	0.030	0.003	2.245***	−2.840	−1.948	−1.508	1.812	2.369	3.236	4.72	393
9	0.028	0.002	2.490**	−2.975	−2.019	−1.553	1.918	2.429	3.361	6.83	390
12	0.027	0.002	2.563**	−3.083	−2.073	−1.565	1.938	2.525	3.595	8.95	387
18	0.022	0.002	2.373***	−3.149	−2.135	−1.654	2.038	2.638	3.755	9.93	381
				Fractiles of Average b							
Avg. b	0.029***			−0.032	−0.023	−0.018	0.024	0.030	0.043		

Table III—*Continued*

K	b	Bias	$Z(b)$	0.01	0.05	0.10	0.90	0.95	0.99	Adj. R^2	N
Panel D: X = VP12(TB)											
				Fractiles of Z-Statistics							
1	0.038	0.002	3.263*	−2.131	−1.466	−1.147	1.454	1.825	2.500	2.58	398
3	0.034	0.002	3.696*	−2.643	−1.772	−1.332	1.779	2.247	3.071	6.21	396
6	0.029	0.002	3.509*	−2.746	−1.839	−1.381	1.844	2.318	3.267	9.22	393
9	0.029	0.002	3.850*	−2.793	−1.855	−1.411	1.915	2.414	3.329	14.06	390
12	0.028	0.002	3.960*	−2.867	−1.911	−1.433	1.978	2.473	3.411	18.21	387
18	0.023	0.002	3.780**	−3.020	−1.960	−1.436	2.102	2.651	3.892	20.63	381
				Fractiles of Average b							
Avg. b	0.030*			−0.021	−0.014	−0.011	0.016	0.020	0.027		
X = Dow Jones Dividend-to-Price Ratio											
				Fractiles of Average b							
Average	0.173	0.031	0.902	−0.367	−0.253	−0.188	0.260	0.334	0.514	1.32	
X = Dow Jones Earnings-to-Price Ratio											
				Fractiles of Average b							
Average	0.036	0.008	0.591	−0.112	−0.077	−0.058	0.075	0.098	0.151	0.25	
X = Dow Jones Book-to-Market Ratio											
				Fractiles of Average b							
Average	0.005	0.001	0.517	−0.017	−0.011	−0.009	0.011	0.015	0.023	0.33	

*, **, *** Significant at the 1-, 5-, and 10-percent levels, respectively.

where X = VP3(TB) or VP3(LT). Results for VP12(TB) and VP12(LT) are similar and are not separately reported. Since V/P is correlated to some extent with DJDP, DJEP, and DJBM, we want to examine whether the predictive power of V/P survives in regressions that include all four variables. Once again, we expect the slope coefficients corresponding to each independent variable to be positive. Therefore, one-sided tests of the null of no predictability are appropriate.

Table IV presents the results of these multivariate forecasting regressions. The columns labeled *p-value* refer to the upper-tail observed significance levels of the corresponding test statistics to the left. These results indicate that only V/P consistently predicts Dow Jones portfolio returns. The Z-statistics corresponding to V/P are significant at the 1 percent or 5 percent level for all horizons for both specifications. DJBM has some predictive power, but is not significant at the 10 percent level. Similarly, Panel A shows that the dividend yield, DJDP, has weak predictive power at long horizons ($K = 18$), but the effect is not significant. Overall, it is clear that only V/P consistently predicts future Dow Jones returns. Similar or stronger results obtain with the S&P 500 and small firm (SFQ1) portfolio returns (not reported). These findings show that the forecasting power of V/P is robust to the inclusion of other intrinsic value measures in the forecasting regression.

B.3. Multivariate Regressions Involving V/P, TB1, Def, and Term

We also examine the forecasting power of V/P controlling for business-cycle-related variation in conditional expected returns. Fama and French (1989) show that the default spread, *Def*, and the term spread, *Term*, predict future stock returns. They interpret these two variables as ex ante measures of default and term risk related to the business cycle. Therefore, they argue that conditional expected stock returns vary with the business cycle. Furthermore, TB1 has predictive power for returns in our sample period (see Table AI) and is an important component in estimating V. Therefore, we test whether V/P continues to predict future stock returns in the presence of TB1. Specifically, we run two separate regressions of the form:

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + b\, Def_t + c\, Term_t + d\, TB1_t + e\, X_t + u_{t+K,t}, \tag{15}$$

where X = VP3(TB) or VP3(LT). Again, replacing VP3(TB) or VP3(LT) with VP12(TB) or VP12(LT) gives similar results.

If the predictive power of these V/P measures comes entirely from their correlation with TB1, *Def*, and *Term*, the slope coefficient on V/P should be insignificant in this regression. Since *Def* and *Term* move countercyclically with the business cycle, we expect high default spread to predict high stock returns and high term spread to predict high stock returns. It is more difficult to predict the direction of the relation between TB1 and future returns. However, if this variable rises and falls with the business cycle, we would expect a negative relation.

Table IV

Multivariate Forecasting Regressions Involving D/P, E/P, B/M, and V/P

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + bDJDP_t + cDJEP_t + dDJBM_t + eX_t + u_{t+K,t}$$

This table summarizes multivariate forecasting regression results using monthly data from May 1963 to June 1996. The dependent variable in these regressions is the excess return on the DJIA portfolio. DJDP refers to the annual dividend yield on the Dow Jones, DJEP refers to the annual earnings yield on the Dow Jones, and DJBM refers to the book-to-market ratio on the Dow Jones. The independent variable X changes in each panel. VPy(TB) refers to the y period value-to-price ratio using 1-month T-bill rates and VPy(LT) refers to the y period value-to-price ratio using long-term Treasury bond rates for the Dow Jones, in percentage. For $K > 1$, the regressions use overlapping observations. b, c, d, and e are slope coefficients from the OLS regressions. $Z(b)$, $Z(c)$, $Z(d)$, and $Z(e)$ are the corresponding asymptotic Z-statistics computed using generalized method of moments (GMM) standard errors with Newey–West correction. These standard errors correct for induced autocorrelation in regression residuals due to overlapping observations and for generalized conditional heteroskedasticity. The columns labeled p-Value refer to the upper tail observed significance levels of the corresponding test statistics to the left. The observed significance levels are obtained by comparing the test statistics to their empirical distribution generated under the null from 5,000 trials of a Monte Carlo simulation. The artificial data for the simulation are generated under the null using the VAR approach. Average represents the average slope statistic. Adj. R^2 refers to the adjusted coefficient of determination from the OLS regression. The column to the right of Adj. R^2 provides the p-value for Adj. R^2 obtained from the same Monte Carlo simulation and can be used to assess overall predictability in the multivariate regression.

K	b	$Z(b)$	p-Value	c	$Z(c)$	p-Value	d	$Z(d)$	p-Value	e	$Z(e)$	p-Value	Adj. R^2	p-Value
Panel A: X = VP3(LT)														
1	0.709	0.86	0.460	−0.102	−0.83	0.205	−0.067	−1.73	0.902	0.116	3.94*	0.000	3.37	0.001
3	0.687	1.05	0.436	−0.041	−0.44	0.122	−0.065	−2.03	0.901	0.098	4.57*	0.000	8.35	0.003
6	0.408	0.68	0.530	−0.030	−0.31	0.149	−0.048	−1.69	0.824	0.087	4.63*	0.001	13.13	0.020
9	0.610	1.13	0.394	−0.065	−0.63	0.286	−0.049	−2.05	0.874	0.079	4.33*	0.003	18.13	0.018
12	0.704	1.40	0.319	−0.072	−0.69	0.343	−0.050	−2.34	0.895	0.074	4.12*	0.005	23.78	0.018
18	0.889	2.15	0.177	−0.114	−1.30	0.592	−0.047	−2.60	0.879	0.060	3.82*	0.010	30.28	0.045
Average	0.668		0.393	−0.071		0.284	−0.054		0.855	0.086***		0.052		
Panel B: X = VP3(TB)														
1	0.214	0.26	0.469	0.099	0.81	0.218	−0.058	−1.50	0.904	0.070	4.60*	0.000	4.39	0.001
3	0.264	0.41	0.442	0.130	1.38	0.136	−0.057	−1.81	0.900	0.060	5.33*	0.000	10.60	0.002
6	0.075	0.13	0.536	0.114	1.27	0.171	−0.040	−1.43	0.824	0.049	4.60*	0.001	14.49	0.019
9	0.309	0.59	0.399	0.062	0.66	0.322	−0.041	−1.78	0.867	0.045	4.20*	0.003	20.63	0.019
12	0.440	0.88	0.319	0.042	0.45	0.380	−0.042	−1.99	0.890	0.041	4.07*	0.005	25.88	0.018
18	0.718	1.65	0.168	−0.035	−0.45	0.625	−0.039	−1.98	0.873	0.030	3.91**	0.012	29.37	0.044
Average	0.337		0.391	0.069		0.307	−0.046		0.871	0.049*		0.001		

*, **, *** Significant at the 1-, 5-, and 10-percent levels, respectively.

Table V presents the results of multivariate forecasting regressions involving TB1, *Def*, and *Term*. TB1 is a significant predictor at the 10 percent level, but its predictive power is much lower than V/P. Interestingly, neither *Def* nor *Term* has much incremental predictive power after controlling for V/P. Recall that *Term* is the difference between the long-term and short-term interest rate, and picks up the slope of the term structure. Table V shows that *Term* has little predictive power in the presence of the other variables, suggesting that the slope of the term structure is not associated with systematic departures of value from price.

Table V shows that V/P continues to forecast the Dow Jones portfolio return, with very little reduction in significance. The *Z*-statistic corresponding to V/P remains significant at the 1 percent level for 1- to 12-month horizons, and is significant at the 5 percent level for the 18-month horizon. The results are even stronger for the S&P 500 and SFQ1 portfolios (not reported). Overall, the evidence shows that V/P predicts future market returns even after controlling for business cycle variation in expected returns.

Table VI presents a further robustness check. The autocorrelations in Table I suggest that there may be mean reversion in large firm stock prices at long horizons. To ensure that our results are not driven by this phenomenon, we reestimate our prediction regression including the 36-month lagged market return. The results, reported in Table VI, show that both VP3(TB) and VP3(LT) continue to predict returns even after controlling for past returns.[29]

C. Alternative Measures of VP

Thus far, the evidence shows that V has superior ability to track variations in the price of the DJIA over time, and V/P has predictive power for subsequent market returns. In this subsection, we examine the incremental contribution of various components of the valuation model to the success of V.

Table VII presents the results of univariate regressions involving alternative measures of V/P. Specifically, we regress excess return for the DJIA separately on three variables: (1) a measure of V/P based on historical earnings forecasts (VHP3(TB)), (2) a measure of V/P based on analysts' forecasts but a constant discount rate of 13 percent (VP3(CR)), and (3) the short-term T-bill rate (TB1). These three regressions can be compared to the result for VP3(TB) reported earlier in Panel B of Table III.

The first set of regressions in Table VII illustrates the effect of replacing analyst forecasts with a time-series model for earnings. We find that although VHP3(TB) continues to have significant predictive power, it is less effective than VP3(TB) across all forecast horizons (compare to Panel B of Table III). This finding demonstrates the incremental usefulness of analyst forecasts once a time-varying discount rate has been incorporated into the valuation model.

[29] Replacing 36-month lagged returns with 12-month or 24-month lagged returns yields similar results. Using 12-period rather than three-period versions of V/P also does not affect these findings.

The second set of regressions shows the effect of replacing time-varying discount rates with a constant discount rate of 13 percent. We find that VP3(CR) has no significant predictive power, demonstrating that a time-varying interest rate component is essential to the success of V/P. Without this component, V/P has no reliable statistical power to predict market returns.

The third set of regressions show that the short-term interest rate alone cannot explain the success of V/P. TB1 has significant predictive power for short-term market returns. Specifically, TB1 is negatively correlated with future market returns in the one- to three-month horizon. However, this predictive power is much weaker than that of V/P at long horizons. Apparently the predictive power of V/P derives, in large part, from the interaction between time-varying interest rates and forecast earnings (or dividends).

Finally, we conclude our analysis by a more extensive evaluation of 25 alternative measures of V/P based on their ability to: (1) track variations in the price of the DJIA over time and (2) predict subsequent DJIA excess returns. This analysis is performed for the subperiod from January 1979 to June 1996. As discussed earlier, this is the subperiod over which we have all the information necessary to estimate the various V/P measures.

The first three variables are DJDP, DJEP, and DJBM, respectively. The other 22 variables are variations of V/P, and are summarized below:

- VHP$y(x)$. These value estimates are computed using time-series historical earnings forecasts for year $t + 1$ to $t + 3$ earnings rather than analysts' forecasts. y represents the number of forecasting periods. x = TB or LT indicates whether the short-term or long-term riskless rate is used, respectively. x = CR indicates the use of a constant 13 percent discount rate.
- VP$y(x)$. Similar to the original V/P, this variable is computed using analysts' forecasts of earnings in the post-1978 period and historical earnings estimates in the pre-1979 period. y represents the number of forecasting periods. x = TB or LT indicates whether the short-term or long-term riskless rate is used, respectively. x = CR indicates that a constant 13 percent discount rate is used. x = TBzF or LTzF indicates these estimates are computed using a z-factor Fama–French (1997) industry risk premium.

The tracking error is a composite measure of the coefficient of variation (standard deviation divided by mean) and the first-order autocorrelation parameter for each value estimate. Lower scores are assigned for lower coefficients of variation and autocorrelation. These two components of the tracking error are scaled to receive approximately equal weight. The predictive ability measure is the average Newey–West adjusted Z-statistic for one-month-ahead and nine-months-ahead Dow Jones returns prediction regressions. Table VIII reports the original components as well as the composite scores for both dimensions.

Table VIII highlights the essential role played by time-varying interest rates. When a constant discount rate is used (rows 16 and 17), V/P has large tracking errors and no significant predictive power. Comparing rows 16 and

Table V

Multivariate Forecasting Regressions Involving Business Cycle Variables

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + bDef_t + cTerm_t + dTBI_t + eX_t + u_{t+K,t}$$

This table summarizes multivariate forecasting regression results using monthly data from May 1963 to June 1996. The dependent variable in these regressions is the excess return on the DJIA portfolios. *Def* is the annualized end-of-month default spread in percentage. *Term* is the annualized end-of-month term spread, and TB1 is the annualized end-of-month yield on the one-month Treasury bill, all in percentage. The independent variable X changes in each panel. VPy(TB) refers to the y period value-to-price ratio using one-month T-bill rates and VPy(LT) refers to the y period value-to-price ratio using long-term T-bond rates, in percentage. For $K > 1$, the regressions use overlapping observations. b, c, d, and e are slope coefficients from the OLS regression. $Z(b)$, $Z(c)$, $Z(d)$, and $Z(e)$ are the corresponding asymptotic Z-statistics computed using generalized method of moments (GMM) standard errors with Newey–West correction. These standard errors correct for induced autocorrelation in regression residuals due to overlapping observations and for generalized conditional heteroskedasticity. The columns labeled p-Value refer to the upper tail observed significance levels of the corresponding test statistics to the left. The observed significance levels are obtained by comparing the test statistics to their empirical distribution generated under the null from 5,000 trials of a Monte Carlo simulation. The artificial data for the simulation are generated under the null using the VAR approach. Average represents the average slope statistic. Adj. R^2 refers to the adjusted coefficient of determination from the OLS regression. The column to the right of Adj. R^2 provides the p-value for Adj. R^2 obtained from the same Monte Carlo simulation and can be used to assess overall predictability in the multivariate regression.

K	b	$Z(b)$	p-Value	c	$Z(c)$	p-Value	d	$Z(d)$	p-Value	e	$Z(e)$	p-Value	Adj. R^2	p-Value
Panel A: X = VP3(LT)														
1	3.337	1.61	0.230	−0.127	−0.61	0.753	−0.541	−4.09***	0.932	0.094	2.98*	0.000	7.30	0.003
3	3.568	2.49	0.224	−0.128	−0.86	0.589	−0.454	−4.95***	0.913	0.079	3.35*	0.001	16.74	0.010
6	3.268	2.68	0.323	−0.132	−1.06	0.538	−0.369	−5.16	0.853	0.069	3.88*	0.002	24.34	0.020
9	2.347	2.08	0.211	0.012	0.12	0.629	−0.264	−4.25***	0.900	0.055	3.93*	0.003	28.78	0.022
12	2.089	1.99	0.159	0.041	0.44	0.639	−0.212	−3.59***	0.919	0.048	4.01*	0.005	32.65	0.015
18	2.279	2.53***	0.069	0.018	0.18	0.783	−0.134	−2.25***	0.927	0.032	3.35**	0.016	33.13	0.019
Average	2.815		0.337	−0.053		0.589	−0.329		0.580	0.063*		0.000		
Panel B: X = VP3(TB)														
1	3.035	1.45	0.224	−0.503	−1.69	0.730	−0.524	−4.10***	0.935	0.073	2.90*	0.001	7.06	0.003
3	3.322	2.23	0.212	−0.427	−1.93	0.562	−0.434	−4.76***	0.926	0.059	3.07*	0.001	15.84	0.008
6	3.063	2.36	0.311	−0.364	−2.08	0.515	−0.341	−4.61	0.862	0.049	3.28*	0.002	21.90	0.019
9	2.195	1.86	0.205	−0.167	−1.31	0.602	−0.239	−3.69***	0.906	0.039	3.26*	0.005	25.93	0.021
12	1.969	1.83	0.159	−0.115	−1.02	0.606	−0.189	−3.08***	0.927	0.034	3.32*	0.009	29.63	0.015
18	2.211	2.48***	0.062	−0.097	−0.86	0.757	−0.123	−2.11***	0.934	0.024	2.89**	0.022	31.50	0.020
Average	2.633		0.317	−0.279		0.524	−0.308		0.575	0.046*		0.000		

*, **, *** Significant at the 1-, 5-, and 10-percent levels, respectively.

Table VI

Multivariate Forecasting Regressions Including Lagged Returns

$$\sum_{k=1}^{K} DJ_{t+k}/K = a + bX_t + c\sum_{l=1}^{36} DJ_{t+1-l}/36 + u_{t+K,t}$$

This table summarizes multivariate forecasting regression results involving the sum of lagged DJIA returns. The dependent variable in these regressions is the excess return on the DJIA portfolio. The independent variable X changes in each panel. VPy(TB) refers to the y period value-to-price ratio using one-month T-bill rates and VPy(LT) refers to the y period value-to-price ratio using long-term T-bond rates, in percentages. For $K > 1$, the regressions use overlapping observations. b and c are the slope coefficients from the OLS regression. $Z(b)$ and $Z(c)$ are the asymptotic Z-statistics computed using generalized method of moments (GMM) standard errors with Newey–West correction. These standard errors correct for induced autocorrelation in regression residuals due to overlapping observations and for generalized conditional heteroskedasticity. Adj. R^2 refers to the adjusted coefficient of determination from the OLS regression. N is the number of monthly observations.

	X = VP3(TB)						X = VP3(LT)					
K	b	$Z(b)$	c	$Z(c)$	Adj. R^2	N	b	$Z(b)$	c	$Z(c)$	Adj. R^2	N
1	0.056	4.04*	−0.022	−0.04	4.20	362	0.064	2.82*	0.082	0.15	2.01	362
3	0.048	4.71*	−0.104	−0.22	9.68	360	0.056	3.20*	−0.004	−0.01	5.22	360
6	0.040	4.32*	−0.171	−0.45	13.95	357	0.048	3.25*	−0.059	−0.15	8.72	357
9	0.037	4.33*	−0.148	−0.51	19.67	354	0.043	3.22**	−0.050	−0.16	11.15	354
12	0.035	4.29*	−0.146	−0.61	24.54	351	0.041	3.11**	−0.042	−0.15	14.21	351
18	0.029	4.07*	−0.010	−0.05	25.27	345	0.035	2.92**	0.092	0.38	15.44	345

*, ** Significant at the 1- and 5-percent levels, respectively.

Table VII

Univariate Forecasting Regressions Involving Alternate Measures of V/P

$$\sum_{k=1}^{K} Y_{t+k}/K = a + bX_t + u_{t+K,t}$$

This table summarizes univariate forecasting regression results using monthly data from May 1963 to June 1996. For $K > 1$, the regressions use overlapping observations. b is the slope coefficient from the OLS regression. $Z(b)$ is the asymptotic Z-statistic computed using generalized method of moments (GMM) standard errors with Newey–West correction. These standard errors correct for induced autocorrelation in regression residuals due to overlapping observations and for generalized conditional heteroskedasticity. Adj. R^2 refers to the adjusted coefficient of determination from the OLS regression, in percentage. The label X refers to the explanatory variable and the label Y refers to the dependent variable. The notation VHP refers to value-to-price ratios computed using historical earnings. The notation VP3(CR) refers to value-to-price ratios computed using constant discount rates. TB1 is the annualized end-of-month yield on the one-month Treasury bill, in percentage.

	Y = DJ, X = VHP3(TB)			Y = DJ, X = VP3(CR)			Y = DJ, X = TB1			
K	b	$Z(b)$	Adj. R^2	b	$Z(b)$	Adj. R^2	b	$Z(b)$	Adj. R^2	N
1	0.038	2.72*	2.20	0.007	0.76	−0.09	−0.227	−3.01*	1.49	398
3	0.034	2.96**	5.10	0.008	0.96	0.30	−0.171	−2.85*	2.57	396
6	0.029	2.86**	7.54	0.009	1.12	1.20	−0.116	−1.85	2.38	393
9	0.026	2.82**	9.46	0.011	1.42	2.96	−0.090	−1.33	2.26	390
12	0.023	2.65**	10.62	0.012	1.67	4.89	−0.065	−0.98	1.57	387
18	0.016	2.06***	8.35	0.011	1.98	7.57	−0.016	−0.31	−0.08	381

*, **, *** Significant at the 1-, 5-, and 10-percent levels, respectively.

Table VIII

Tracking Error and Predictive Ability of Alternative Value Measures (January 1979 to June 1996)

This table presents a comparison of alternative value estimates based on their ability to track variations in the price of the DJIA over time ("tracking error") and predict subsequent DJIA excess returns ("predictive ability"). The sample period, January 1979 to June 1996, is the period for which information is available to estimate all these alternative measures. The predictive ability measure is the average Newey–West adjusted Z-statistic for one-month-ahead and nine-month-ahead Dow Jones returns prediction regressions. The tracking error is a composite measure of the coefficient of variation and first-order autocorrelation for each price/value ratio estimate. Lower scores are assigned for lower coefficients of variation and autocorrelation. The two components are scaled to receive approximately equal weight. VPy represents value-to-price where value is estimated using a residual income model with y forecasting periods. Descriptions of the discount rates used are in parentheses. TB represents short-term T-bill and LT represents a long-term T-bond rate. zF indicates the use of a z-factor Fama and French (1997) industry risk premium rather than a marketwide equity risk premium. CR represents the use of a 13 percent constant discount rate for the entire time period.

Variable	Mean	Std. Dev.	AR1 Parameter	Composite Tracking Error	Predictive Z Next 1-Month	Predictive Z Next 9-Month	Composite Predictive Ability
1 DJDP	3.98	1.18	0.97	2.24	−0.45	−0.15	−0.30
2 DJEP	7.49	3.58	0.97	2.62	−0.22	−0.18	−0.20
3 DJBM	65.46	24.43	0.98	2.49	−0.50	−0.14	−0.32
4 VP3(TB)	76.59	14.50	0.85	0.99	2.30	3.82	3.06
5 VP12(TB)	80.34	16.08	0.85	1.01	2.16	3.31	2.73
6 VP18(TB)	91.79	23.02	0.78	0.52	2.65	3.03	2.84
7 VP3(LT)	62.40	9.23	0.91	1.42	1.24	1.71	1.47
8 VP12(LT)	65.26	11.89	0.94	1.75	0.82	1.05	0.94
9 VP18(LT)	69.84	11.62	0.84	0.86	1.57	2.46	2.02

10 VHP3(TB)	65.20	14.14	0.90	1.48	1.98	2.52	2.25
11 VHP12(TB)	70.40	14.77	0.89	1.38	1.84	2.35	2.10
12 VHP18(TB)	71.88	13.66	0.83	0.82	2.48	4.01	3.24
13 VHP3(LT)	53.61	11.68	0.96	1.99	0.77	0.63	0.70
14 VHP12(LT)	57.79	13.31	0.96	2.02	0.56	0.48	0.52
15 VHP18(LT)	56.53	11.31	0.95	1.87	0.97	0.80	0.88
16 VP3(CR)	88.38	27.58	0.97	2.27	−0.37	0.12	−0.12
17 VHP3(CR)	76.02	27.78	0.98	2.47	−0.42	−0.13	−0.28
18 VP3(TB3F)	101.80	26.03	0.84	1.04	2.33	3.66	3.00
19 VP12(TB3F)	118.66	39.91	0.83	1.13	3.04	2.60	2.82
20 VP3(TB1F)	104.28	26.59	0.86	1.21	2.72	3.57	3.15
21 VP12(TB1F)	116.09	34.04	0.83	1.04	2.67	4.06	3.37
22 VP3(LT3F)	74.94	11.57	0.90	1.35	1.90	2.07	1.98
23 VP12(LT3F)	81.28	14.62	0.90	1.40	1.89	1.85	1.87
24 VP3(LT1F)	77.31	11.97	0.90	1.35	1.66	1.72	1.69
25 VP12(LT1F)	80.83	12.89	0.91	1.44	1.58	1.67	1.63

17, it is clear that the inclusion of analyst forecasts does little to improve performance in the absence of time-varying interest rates. Rows 13 to 15 show that the performance of V/P on both dimensions improves with the addition of long-term interest rates. Further improvement is seen with the addition of analyst forecasts (rows 7 to 9). Apparently the inclusion of analyst forecasts enhances the performance of V/P when time-varying interest rates are already used.

The biggest performance improvement is obtained with the addition of short-term interest rates. Both the tracking ability and predictive power of V/P increase sharply when short-term interest rates are used (rows 10 to 12). Comparing rows 10 to 12 with rows 4 to 6, we see once again that analyst forecasts are incrementally useful when time-varying short-term interest rates are used to compute V. For example, VP3(TB) (row 4) has both lower tracking error and higher predictive power than VHP3(TB) (row 10).

Rows 18 to 25 show that adding the Fama and French (1997) industry risk premia does little to improve overall tracking ability or predictive power. Variations in the forecast horizon show that the 18-period versions of V/P seem to have an advantage over the 3- or 12-period versions. But this improvement is not consistent throughout the range of discount rates and earnings forecasting methods.

The information contained in Table VIII is graphically illustrated in Figure 4. This figure plots the composite tracking error on the horizontal axis and the predictive ability on the vertical axis. To highlight the benefit of using time-varying interest rates and analyst forecasts, the individual observations are represented by symbols that differ depending on the discount rate and the type of earnings forecast method used.

Several interesting observations appear in this graph. First, there is a strong negative relation between tracking error and predictive ability. Empirical value estimates that track prices better over time also tend to have greater predictive power for subsequent returns. Second, the performance of the V/P measure improves with the addition of time-varying interest rates and analyst forecasts. This is evident in the clustering of observations with a common symbol. The number of forecast periods and the choice of the risk premium are of only secondary importance.

The traditional DJDP, DJBM, and DJEP measures appear at the lower right corner of the graph, along with V/P estimates that do not feature time-varying interest rates. Observations represented by the diamond-shaped symbols show that the use of long-term bond rates improves both tracking and predictive power. The solid squares show that the addition of analyst forecasts further reduces tracking error, and also enhances the predictive power of the V estimates. The open and solid circles show that using short-term rather than long-term interest rates again adds to the performance of V/P. Introduction of the Fama–French risk premia and variations in the forecast horizon generally added little to V's performance.

The upper left corner of this graph depicts estimates of V that minimize tracking error while also maximizing predictive power. Three version of V are on the "efficient frontier," as indicated by the dashed line segments. These three

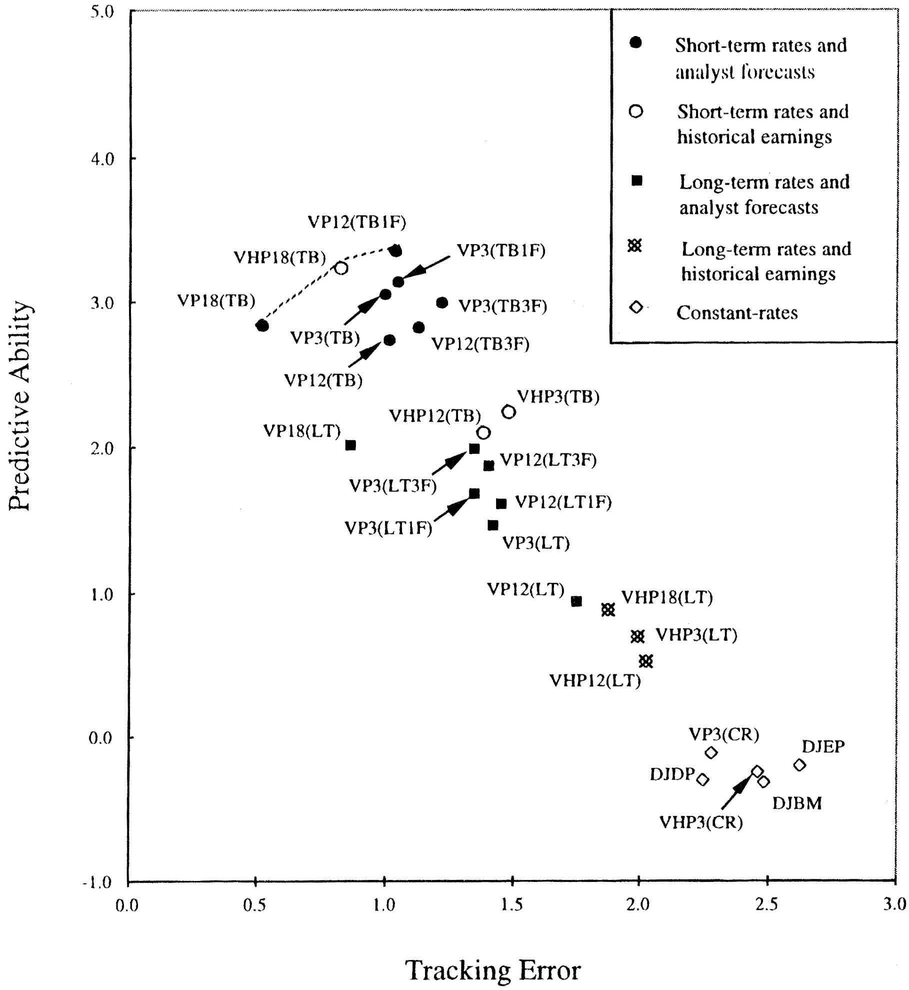


Figure 4. A comparison of alternative intrinsic value estimates in terms of their predictive ability and tracking error. This graph plots the predictive ability and tracking error for alternative intrinsic value estimates during the period January 1979 to June 1996. The value estimates are as described in Table VIII. The dashed line depicts an "efficient frontier" for minimum tracking error and maximum predictive power. The predictive ability measure depicted is the average Newey–West adjusted Z-statistic for one-month-ahead and nine-months-ahead Dow Jones returns prediction regressions. The tracking error is a composite measure of the coefficient of variation and first-order autocorrelation for each value-to-price ratio estimate. Lower scores are assigned for lower coefficients of variation and autocorrelation. The two components of the tracking error calculation are scaled to receive approximately equal weighting.

estimates are not dominated by the others and are, by our benchmarks, the best estimates of the intrinsic value of the DJIA. Note that the two value metrics we focus on throughout this paper (VP3(TB) and VP3(LT)) are not on the efficient frontier when placed in the context of these 25 measures.

VI. Summary

The purpose of this paper is to develop measures of the intrinsic value for the DJIA independent of its market price. We model the time-series relation between price and value as a cointegrated system rather than a static equality. In this context, we examine the relative performance of alternative proxies for the market's intrinsic value in terms of: (a) their ability to track movements in the index price over time and (b) the ability of the alternative V/P ratios to predict subsequent market returns.

It is important for the reader to understand the various levels at which our analysis is useful. First, we contribute to the literature on stock valuation by modeling the time-series relation between price and value as a cointegrated system. This approach leads to a two-dimensional framework for evaluating the "usefulness" of an intrinsic value estimate, with significant implications for market-related research in both finance and accounting.

Traditionally, in accounting literature the "value-relevance" of a fundamental signal is measured in terms of the strength of its correlation with contemporary returns. Signals that track current returns better (worse) are deemed to reflect "good" ("bad") accounting. We show that when price is a noisy measure of value, the value-relevance of a fundamental signal can also be evaluated in terms of its ability to contribute to return prediction. Under reasonable assumptions, superior value estimates produce V/P ratios that better predict returns.

Whether one dimension is more important than the other depends on the decision context. For example, portfolio managers may be more interested in predictive power, and accounting regulators may be more interested in tracking ability. In either case, the methodology introduced in this paper allows researchers to simultaneously assess the usefulness of an intrinsic value estimate along both dimensions.

Second, our study contributes to the literature on the predictability of market returns. Specifically, we show that in the 1963–1996 period, traditional value benchmarks such as B/P, E/P, and D/P have little predictive power for overall returns in the United States. Using a richer valuation model, we develop a V measure that outperforms these simple market multiples in terms of both tracking ability and predictive power.

Our evidence shows that market returns over the 1963–1996 time period are predictable on the basis of a more robust measure of intrinsic value. This predictability is not due to the general mean-reversion pattern observed in postwar U.S. market returns, nor is it due to known term-structure-related variables, or other traditional price-to-value indicators. Although our finding is consistent with market inefficiency, we cannot rule out the possibility that the predictive power of V/P arises from time-varying expected returns. Our tests control for all known determinants of such risk. However, it is still possible that V/P captures a new dimension of time-varying risk that has not yet been identified.

Setting aside the market efficiency debate, our results provide important insights into the components of the residual income valuation model in time-series applications. Specifically, we show that a primary contributor to the success of V is the inclusion of time-varying interest rates. Intrinsic value measures that incorporate a time-varying interest rate component produce V/P ratios with much better tracking ability and predictive power. Intrinsic value estimates that omit time-varying interest rates have little predictive power for future returns.

Both long-term Treasury bond rates and short-term T-Bill rates are useful; however, we find that V estimates based on the short-term interest rate outperform those based on the long-term rate. Our result implies that better empirical estimates of intrinsic value may be derived by using short-term rates. However, given our limited sample, we recommend caution in generalizing this result to other firms and time periods.

Furthermore, we find that using mean I/B/E/S analyst forecasts, rather than forecasts based on a time-series of historical earnings, improves both the tracking ability of V and the predictive power of V/P. When used alone, analyst earnings forecasts have little predictive power for returns. However, when used in conjunction with time-varying discount rates, these consensus forecasts can enhance the performance of V estimates. The shortcomings and biases of sell-side analyst earnings forecasts are well known. Yet we show that in time-series applications, analyst earnings forecasts can help produce better intrinsic value estimates.

Finally, our results show that V/P's ability to predict marketwide returns derives from the interaction between time-varying discount rates and expected future earnings. Despite their importance in estimating V, time-varying interest rates do not fully explain the success of V/P. V/P and short-term T-bill rates are only correlated at the 0.26 to 0.30 level. Moreover, V/P continues to have strong predictive power when we include the short-term interest rate, the term structure risk premium, and the default risk premium in the prediction regression. This evidence suggests that if V/P is a risk proxy, the nature of this risk is not readily captured by existing interest rate proxies. We leave the exact explanation for the predictive power of V/P to future research.

Appendix. Monte Carlo Simulation Methodology

The simulation methodology we use closely follows that used by Hodrick (1992) and Swaminathan (1996). Define $Z_t = (\mathrm{DJ}_t, \mathrm{SP500}_t, \mathrm{SFQ1}_t, \mathrm{DJDP}_t, \mathrm{DJEP}_t, \mathrm{DJBM}_t, \mathrm{VP}_t, \mathit{Def}_t, \mathit{Term}_t, \mathrm{TB1}_t)'$ where Z_t is a 10×1 column vector. We fit a first-order VAR to Z_t using the following specification:

$$Z_{t+1} = A_0 + A_1 Z_t + u_{t+1}, \tag{A1}$$

where A_0 is a 10×1 vector of intercepts and A_1 is a 10×10 matrix of VAR coefficients, and u_{t+1} is a 10×1 vector of VAR residuals. The VAR results are presented in Table AI. The estimated VAR is used as the data generating process (DGP) for the simulation.

The point estimates in Table AI are used to generate artificial data for the Monte Carlo simulations. We first impose the null hypothesis of no predictability on the three returns, DJ, SP500, and SFQ1, in the VAR. This is done

Table AI
First-Order Vector Autoregressions

This table summarizes the result of the first-order vector autoregressions, computed using monthly data from May 1963 to June 1996. DJ, SP500, and SFQ1 are continuously compounded excess returns of the Dow Jones stock portfolio, S&P 500 stock portfolio, and the smallest size quintile of NYSE stocks respectively, expressed in percentage. DJDP refers to the annual dividend yield on the Dow Jones, DJEP refers to the annual earnings yield on the Dow Jones, and DJBM refers to the book-to-market ratio on the Dow Jones. VP3TB refers to the value-to-price ratio for the Dow Jones using a three-period model and one-month T-bill rates. *Def* is the annualized end-of-month default spread in percentage. *Term* is the annualized end-of-month term spread, and TB1 is the annualized end-of-month yield on the one-month Treasury bill, all in percentage. The first column refers to the left-hand side variables in the VAR and columns 3–12 refer to the slope coefficients of the right-hand side variables. Chisq(10) is the chi-square test statistic with nine degrees of freedom testing the null hypothesis that the slope coefficients are jointly zero. Adj. R^2 refers to the adjusted coefficient of determination, in percentage. The numbers in parentheses are the asymptotic Z-statistics computed using the White heteroskedasticity consistent standard errors.

Dep.	Intpt.	DJ(t)	SP500(t)	SFQ1(t)	DJDP(t)	DJEP(t)	DJBM(t)	VP3TB(t)	Def(t)	Term(t)	TB1(t)	Chisq(10)	Adj. R^2
DJ(t+1)	−2.505	0.144	−0.154	0.014	1.577	0.153	−0.084	0.074	1.845	−0.492	−0.606	40.04	7.07
	(−2.45)	(0.90)	(−0.92)	(0.26)	(1.64)	(1.12)	(−2.07)	(2.69)	(0.80)	(−1.57)	(−4.12)		
SP500(t+1)	−2.545	0.091	−0.131	0.037	1.878	0.209	−0.094	0.056	0.842	−0.264	−0.554	44.47	6.94
	(−2.79)	(0.62)	(−0.83)	(0.67)	(2.14)	(1.64)	(−2.60)	(2.12)	(0.39)	(−0.90)	(−4.08)		
SFQ1(t+1)	−3.833	0.386	−0.270	0.074	3.373	0.299	−0.129	0.077	2.373	−0.920	−1.099	58.19	11.25
	(−2.86)	(1.87)	(−1.30)	(0.87)	(2.59)	(1.59)	(−2.542)	(2.004)	(0.818)	(−2.054)	(−5.339)		
DJDP(t+1)	0.186	−0.001	−0.002	0.001	0.862	−0.008	0.005	−0.002	−0.007	0.005	0.030	7641.86	96.14
	(3.580)	(−0.076)	(−0.252)	(0.315)	(17.964)	(−1.145)	(2.906)	(−1.463)	(−0.072)	(0.295)	(3.693)		
DJEP(t+1)	0.306	−0.025	0.029	0.000	−0.201	0.907	0.011	0.004	−0.301	−0.071	0.075	7519.79	95.70
	(2.276)	(−0.978)	(0.902)	(−0.030)	(−1.581)	(27.611)	(2.329)	(0.839)	(−1.273)	(−1.170)	(2.925)		
DJBM(t+1)	2.532	0.012	−0.057	0.044	−1.144	−0.190	1.048	−0.043	−1.119	0.170	0.514	10121.46	97.26
	(2.732)	(0.080)	(−0.320)	(0.854)	(−1.406)	(−1.632)	(33.481)	(−1.813)	(−0.669)	(0.566)	(3.587)		
VP3TB(t+1)	2.655	0.128	−0.443	0.049	−1.714	−0.487	0.112	0.926	6.162	−0.476	0.627	2036.09	88.04
	(1.663)	(0.588)	(−1.891)	(0.568)	(−1.280)	(−1.842)	(2.082)	(19.378)	(2.043)	(−0.853)	(2.575)		
Def(t+1)	0.002	−0.006	0.006	0.000	0.023	−0.002	−0.001	0.000	0.861	0.003	0.005	2575.49	83.77
	(0.133)	(−2.652)	(2.472)	(−0.246)	(1.544)	(−0.622)	(−1.513)	(−0.080)	(26.621)	(0.650)	(2.038)		
Term(t+1)	−0.103	−0.017	0.021	−0.008	−0.059	−0.077	0.005	0.008	1.172	0.733	0.008	965.14	77.56
	(−0.583)	(−0.683)	(0.705)	(−0.738)	(−0.443)	(−2.981)	(0.950)	(1.786)	(4.062)	(12.682)	(0.291)		
TB1(t+1)	0.281	0.027	−0.027	0.011	0.014	0.073	−0.003	−0.009	−1.110	0.232	0.991	2096.05	91.08
	(1.406)	(0.863)	(−0.783)	(0.854)	(0.093)	(2.380)	(−0.468)	(−1.770)	(−3.858)	(3.366)	(32.892)		

by setting the slope coefficients on the explanatory variables to zero for all of the returns and by setting the intercepts equal to the unconditional means. We use the fitted VAR under the null hypothesis of no predictability to generate 399 observations of the state variable vector, (DJ_t, $SP500_t$, $SFQ1_t$, $DJDP_t$, $DJEP_t$, $DJBM_t$, V/P_t, Def_t, $Term_t$, $TB1_t$). The initial observation for this vector is drawn from a multivariate normal distribution with mean equal to the historical mean and variance-covariance matrix equal to the historical estimated variance-covariance matrix of the vector of state variables.

Once the VAR is initiated, shocks for subsequent observations are generated by randomizing (sampling without replacement; see Noreen (1989)) among the actual VAR residuals. The VAR residuals for DJ_t, $SP500_t$, $SFQ1_t$, $DJDP_t$, $DJEP_t$, $DJBM_t$, V/P_t, Def_t, $Term_t$, and $TB1_t$ are scaled so that the standard errors computed from these residuals will be equal to the standard errors of DJ_t, $SP500_t$, $SFQ1_t$, $DJDP_t$, $DJEP_t$, $DJBM_t$, V/P_t, Def_t, $Term_t$, and $TB1_t$, respectively. These artificial data are then used to run regressions and generate regression statistics. The process is repeated 5,000 times and empirical distributions of univariate and multivariate regression statistics are obtained. Two FORTRAN numerical recipes subroutines, *ran1* and *gasdev*, are used to generate standard normal random numbers.

REFERENCES

Abarbanell, Jeffrey, and Victor Bernard, 1995, Is the U.S. stock market myopic?, Working paper, University of Michigan.

Bernard, Victor L., 1994, Accounting-based valuation methods, determinants of book-to-market ratios, and implications for financial statement analysis, Working paper, University of Michigan.

Botosan, Christine, 1997, The effect of disclosure level on the cost of equity, *Accounting Review* 72, 323–350.

Campbell, John Y., 1993, Why long horizons? A study of power against persistent alternatives, Working paper, NBER and Princeton University.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Carmel, Jonathan Paul, and Martin R. Young, 1997, Long horizon mean reversion in the stock market: The postwar years, Working paper, University of Michigan.

DeBondt, Werner F. M., and Richard M. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557–581.

Dechow, Patricia, Amy Hutton, and Richard Sloan, 1997, An empirical assessment of the residual income valuation model, Working paper, University of Michigan and Harvard University.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Duffee, Gregory R., 1996, Idiosyncratic variation of Treasury bill yields, *Journal of Finance* 51, 527–551.

Edwards, Edgar O., and Philip W. Bell, 1961, *The Theory and Measurement of Business Income* (University of California Press, Berkeley).

Fairfield, Patricia, Richard J. Sweeney, and Teri L. Yohn, 1996, Accounting classification and the predictive content of earnings, *Accounting Review* 71, 337–356.

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–50.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Feltham, Gerald A., and James A. Ohlson, 1995, Valuation and clean surplus accounting for operating and financial activities, *Contemporary Accounting Research* 11, 689–731.

Frankel, Richard, and Charles M. C. Lee, 1997, Accounting diversity and international valuation, Working paper, University of Michigan and Cornell University.

Frankel, Richard, and Charles M. C. Lee, 1998, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* 25, 283–319.

Hamilton, James T., 1994, *Time Series Analysis* (Princeton University Press, Princeton, N.J.).

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hansen, Lars P., and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Kaplan, Steve, and Richard Ruback, 1995, The valuation of cash flow forecasts: An empirical analysis, *Journal of Finance* 50, 1059–1093.

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

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lee, Charles M.C., 1996, Measuring wealth, *CA Magazine* 129, 32–37.

Lee, Charles M.C., Andrei Shleifer, and Richard H. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75–109.

MacBeth, James D., and David C. Emanuel, 1993, Tactical asset allocation: Pros and cons, *Financial Analysts Journal* 49, 30–43.

Nelson, Charles R., and Myung J. Kim, 1993, Predictable stock returns: The role of small sample bias, *Journal of Finance* 48, 641–661.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Noreen, Eric, 1989, *Computer-Intensive Methods for Testing Hypotheses: An Introduction* (John Wiley and Sons, New York).

Ohlson, James A., 1990, A synthesis of security valuation theory and the role of dividends, cash flows, and earnings, *Contemporary Accounting Research* 6, 648–676.

Ohlson, James A., 1991, The theory of value and earnings, and an introduction to the Ball-Brown analysis, *Contemporary Accounting Research* 7, 1–19.

Ohlson, James A., 1995, Earnings, book values, and dividends in security valuation, *Contemporary Accounting Research* 11, 661–687.

Peasnell, Ken, 1982, Some formal connections between economic values and yields and accounting numbers, *Journal of Business Finance and Accounting* 9, 361–381.

Penman, Stephen H., 1997, A synthesis of equity valuation techniques and the terminal value calculation for the dividend discount model, *Review of Accounting Studies* 2, 303–323.

Penman, Stephen H., and Theodore Sougiannis, 1997, A comparison of dividend, cash flow, and earnings approaches to equity valuation, Working paper, University of California at Berkeley and University of Illinois at Urbana-Champaign.

Perron, Pierre, 1988, Trends and random walks in macroeconomic time series: Further evidence from a new approach, *Journal of Economic Dynamics and Control* 12, 297–332.

Phillips, Peter C. B., 1987, Time series regression with a unit root, *Econometrica* 55, 277–301.

Phillips, Peter C. B., and Pierre Perron, 1988, Testing for a unit root in time series regression, *Biometrika* 75, 335–336.

Preinreich, Gabriel, 1938, Annual survey of economic theory: The theory of depreciation, *Econometrica* 6, 219–241.

Richardson, Matthew, and Tom Smith, 1991, Tests of financial models in the presence of overlapping observations, *Review of Financial Studies* 4, 227–254.

Richardson, Matthew, and James H. Stock, 1989, Drawing inferences from statistics based on multi-year asset returns, *Journal of Financial Economics* 25, 323–348.

Rozeff, Michael, 1984, Dividend yields are equity risk premiums, *Journal of Portfolio Management* 11, 68–75.

Shiller, Robert J., 1984, Stock prices and social dynamics, *The Brooking Papers on Economic Activity* 2, 457–510.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

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

Stewart, G. Bennett, 1991, *The Quest for Value* (Harper-Collins, New York).

Summers, Lawrence, 1986, Does the stock market rationally reflect fundamental values?, *Journal of Finance* 41, 591–602.

Swaminathan, Bhaskaran, 1996, Time-varying expected small firm returns and closed-end fund discounts, *Review of Financial Studies* 9, 845–887.

Wang, J., 1993, A model of intertemporal asset prices under asymmetric information, *Review of Economic Studies* 60, 249–282.