Market Expectations in the Cross Section of Present Values *

Bryan Kelly[†]
University of Chicago
Booth School of Business

Seth Pruitt[‡]
Federal Reserve
Board of Governors

First Version: August 2010
This Version: January 2012

Abstract

Returns and cash flow growth for the aggregate U.S. stock market are highly and robustly predictable. Using information extracted from the cross section of book-to-market ratios, we find an out-of-sample forecasting R^2 as high as 10% for returns 20% for dividend growth at the annual frequency. We present a general economic framework linking aggregate market expectations to disaggregated valuation ratios in a dynamic latent factor system. To derive our forecasts we use a regression-based filter to extract factors driving aggregate expected returns and cash flow growth from the cross section of valuation ratios. Our findings suggest that market discount rates are more volatile and less persistent than previously believed.

*We thank Ken French, Amit Goyal and Robert Shiller for sharing data with us, Gray Calhoun, John Cochrane, George Constantinides, Gene Fama, Rob Engle, Xavier Gabaix, Jim Hamilton, Ralph Koijen, Juhani Linnainmaa, Hanno Lustig, Toby Moskowitz, Lubos Pastor, Stijn Van Nieuwerburgh and seminar participants at Barclays Capital, Chicago (Booth), Duke (Economics), the Federal Reserve Board, Michigan (Ross), Northwestern (Kellogg), UCSD (Rady), UNC (Kenan-Flagler) and Yale (SOM) for helpful comments.

[†]Mail: 5807 S. Woodlawn Ave., Chicago, IL 60637. E-mail: bryan.kelly@chicagobooth.edu. Phone: (773) 702-8359.

[‡]Mail: Board of Governors of the Federal Reserve System, Washington, DC 20551. E-mail: seth.j.pruitt@frb.gov. Phone: (202) 452-2344. This paper reflects the views and opinions of the authors and does not reflect on the views or opinions of the Federal Reserve Board, the Federal Reserve System, or their respective staffs.

Explaining market price behavior of the U.S. capital stock is among the most fundamental challenges facing economists. The present value relationship between prices, discount rates and future cash flows has proved a valuable lens for understanding stock price variation. It reveals that price changes are wholly driven by fluctuations in investors' expectations of future returns and cash flow growth. Understanding asset prices amounts to understanding the dynamic behavior of these expectations.

The most common approach to measuring aggregate return and cash flow expectations is predictive regression. As suggested by the present value relation, research has found the aggregate book-to-market ratio to be among the most informative predictive variables. Typical in-sample estimates find that about 10% of annual return variation (1% of annual dividend growth variation) can be explained by forecasts based on the aggregate book-to-market ratio, but that out-of-sample the predictive power vanishes.[1] In this paper we find that reliance on aggregate quantities drastically understates the degree of value ratios' predictive content for both returns and cash flow growth, and hence understates the volatility and accuracy of investor expectations. Our estimates suggest that as much as 10% of the *out-of-sample* variation in annual market returns (25% for dividend growth), and somewhat more of the in-sample variation, can be explained by the cross section of past disaggregated value ratios.[2]

To harness the disaggregated information we represent the cross section of asset-specific book-to-market ratios as a dynamic latent factor model. We relate these disaggregated value ratios to aggregate expected market returns and cash flow growth. Our model highlights the idea that the same dynamic state variables driving aggregate expectations also govern the dynamics of the entire panel of asset-specific valuation ratios. This representation allows us to exploit rich cross-sectional information to extract precise estimates of the market's

[1] See Cochrane (2005) and Koijen and Van Nieuwerburgh (2010) for surveys of return and cash flow predictability evidence using the aggregate price-dividend ratio. Similar results obtain from forecasts based on the aggregate book-to-market ratio.

[2] This out-of-sample performance is robust across many choices of training sample period (see Figures 3-5 below), addressing concerns about out-of-sample statistics recently discussed by Hansen and Timmermann (2011) and Inoue and Rossi (2011) among others.

expectations.

Our approach attacks a challenging problem in empirical asset pricing: How does one exploit a wealth of predictors in relatively short time series? If the predictors number near or more than the number of observations, the standard ordinary least squares (OLS) forecaster is known to be poorly behaved or nonexistent (see Huber (1973)). Our solution is to use partial least squares (PLS, Wold (1975)), which is a simple regression-based procedure designed to parsimoniously forecast a single time series using a large panel of predictors. We use it to construct a *univariate* forecaster for market returns (or cash flow growth) that is a linear combination of assets' valuation ratios. The weight of each asset in this linear combination is based on the covariance of its value ratio with the forecast target. Much of our analysis relies on results from Kelly and Pruitt (2011), who derive the properties of PLS in the factor model setting that applies directly to the asset pricing model considered here.

Using data from 1930-2009, PLS forecasts based on the cross section of portfolio-level book-to-market ratios achieve an out-of-sample predictive R^2 as high as 9.9% for annual market returns, 25.4% for annual dividend growth and 0.8% for monthly market returns (in-sample R^2 of 32.9%, 44.5% and 4.1%, respectively). Since we construct a single factor from the cross section, our results can be compared directly with univariate forecasts from the many alternative predictors that have been considered in the literature. In contrast to our results here, previously studied predictors typically perform well in-sample but become insignificant out-of-sample, often performing worse than forecasts based on the historical mean return (Goyal and Welch (2008)).

Our estimates shed new light on the dynamic processes for expected one-year-ahead returns and cash flow growth rates. We find that the volatility of expected one-year returns since 1955 is 6.9% based on out-of-sample estimates (6.7% from in-sample estimates), nearly 80% higher than the volatility estimated from the aggregate book-to-market ratio. We also find much less persistence in expected returns, with an autocorrelation of 0.20 (0.44 from in-sample estimates), contrasting the persistence of 0.91 based on the aggregate book-to-

market ratio. The evidence for expected market cash flow growth is similar. This degree of variability in short term expectations is difficult to reconcile with fundamentals-based structural asset pricing models, which generate persistent, low volatility fluctuations in expected market returns (e.g. Campbell and Cochrane (1999) or Bansal and Yaron (2004)).

We establish the robustness of our main findings in a number of ways. We evaluate various degrees of portfolio aggregation and find similar results whether we use six, 25 or 100 size and book-to-market sorted portfolios, with forecast performance that increases in the number of portfolios. Applying our method to individual stocks rather than portfolios corroborates our main results. Using the entire cross section of CRSP stocks (a cross section of several thousand value ratios), we find an out-of-sample one month return forecasting R^2 of nearly 1.6%. Sensitivity analysis of out-of-sample predictive performance to different subsamples shows that our results are robust to virtually any choice of sample split after 1955. Our results are also robust to data from outside the U.S. We forecast returns on the value-weighted aggregate world portfolio (excluding the U.S.) by applying PLS to an international cross section of non-U.S., country-level valuation ratios, available beginning in 1975. We find an out-of-sample predictive R^2 of 2.3% (5.3% in-sample) at the monthly frequency, corroborating our results in U.S. data. Lastly, we find similar results when applying PLS to the cross section of portfolio price-dividend ratios rather than book-to-market ratios.

Why do disaggregated prices produce such accurate forecasts? To illustrate the advantages of cross section information, consider a simple CAPM example.[3] In particular, suppose one period expected market returns μ_t and expected return on equity g_t are the two common factors in the economy, and the book-to-market ratio of any asset i is

$$v_{i,t} = a_i - b_{i,\mu}\mu_t + b_{i,g}g_t + e_{i,t} \tag{1}$$

[3]The present value system in Equations 1 and 2 obtains as a special case of the model in Section I. It arises in an economy where μ_t and g_t each follow an AR(1), individual expected returns obey an exact one factor model as in the CAPM, $\mu_{i,t} = \mu_{i,0} + c_{i,\mu}\mu_t$, and individual expected return on equity obeys a one factor model, $g_{i,t} = g_{i,0} + c_{i,g}g_t + e_{i,t}$. This special case is similar to the formulation of Polk, Thompson and Vuolteenaho (2006).

while the market book-to-market ratio is

$$v_t = a - b_\mu \mu_t + b_g g_t. \tag{2}$$

Equation 2 highlights the predictive relationship between v_t, realized market returns ($r_{t+1} = \mu_t + \epsilon^r_{t+1}$) and return on equity ($\Delta cf_{t+1} = g_t + \epsilon^d_{t+1}$). However, it also evokes limitations of the aggregate system for understanding market expectations. Predictive regressions of r_{t+1} on v_t take the form

$$\mathbb{E}_t[r_{t+1}|v_t] = \hat{a} + \hat{b}v_t = \hat{a} + \hat{b}(b_\mu \mu_t + b_g g_t)$$

and thus are unduly influenced by information about return on equity. The reciprocal problem arises in forecasting Δcf_{t+1}. To overcome this difficulty, researchers have taken present value approaches that account for the joint relationship among v_t, μ_t and g_t (see Cochrane 2008b, Lettau and Van Nieuwerburgh 2008, van Binsbergen and Koijen 2010). While this begins to disentangle the link between prices and expectations, these joint systems continue to rely solely on aggregate variables. Because both μ_t and g_t are latent, each adds noise to the signal extraction problem of the other.[4] If there exist other signals for μ_t and g_t in the economy, incorporating them will improve estimates of the latent expectations. This is how disaggregated valuation ratios in Equation 1 become a valuable information source as long as each $v_{i,t}$ provides a non-redundant signal for μ_t and g_t.

PLS conveniently reduces the many available signals to an optimal forecast with a series of ordinary least squares regressions. The first stage consists of "reverse" regressions in which individual valuation ratios for each asset are regressed on the forecast target. Next, in each time period, we run second stage cross-sectional regressions of assets' valuation ratios on firm-specific regression coefficients estimated in the first stage. In the final stage, aggregate return and cash flow growth realizations are regressed on the fitted factors from the second stage, delivering our final filtered estimates for unobservable return and growth expectations.

[4]This remains true despite the absence of measurement error in the aggregate book-to-market expression, as pointed out by Fama and French (1988).

The final-stage predictor is a discerningly-constructed linear combination of disaggregated value ratios that parsimoniously incorporates information from individual valuation ratios into predictions of future aggregate returns and cash flow growth.

The preceding CAPM example is also useful to develop intuition for how PLS works. Each $v_{i,t}$ is a function of only the *expected* portion of returns and cash flows and is uncorrelated with their unanticipated future shocks. Therefore, first stage time series regression coefficients of $v_{i,t}$ on r_{t+1} and Δcf_{t+1} (which serve as observable proxies for the latent factors μ_t and g_t) describe how each firm's valuation ratio depends on the true factors μ_t and g_t. When the coefficients $b_{i,\mu}$ and $b_{i,g}$ differ across i, fluctuations in μ_t and g_t cause the cross section of value ratios to fan out and compress over time. The first-stage coefficient estimates provide a map from the cross-sectional distribution of v_i's to the latent factors. Second-stage cross section regressions of $v_{i,t}$ on first-stage coefficients use this map to estimate the factors at each point in time. Because the first-stage regression takes an errors-in-variables form, second-stage regressions estimate latent expectations $(\mu_t, g_t)'$ with a multiplicative bias. Since OLS forecasts are invariant to scalar multiples of regressors, the third-stage regression of realized returns and growth on the estimated factors delivers consistent estimates of μ_t and g_t.

The earliest predecessors of this study use predictive regressions based on the aggregate price-dividend ratio, including Rozeff (1984), Campbell and Shiller (1988), Fama and French (1988). A large subsequent literature has evolved testing the predictive relation based on aggregate valuation ratios, including Stambaugh (1986, 1999), Hodrick (1992), Goetzmann and Jorion (1993), Nelson and Kim (1993), Kothari and Shanken (1997), Lewellen (2004), Paye and Timmermann (2006), and Pástor and Stambaugh (2009a).

More directly, our paper builds on recent literature that exploits the present value relation to identify market expectations for returns and dividends, including van Binsbergen and Koijen (2010), Ghosh and Constantinides (2010), Ferreira and Santa-Clara (2010), Cochrane (2008a,b), Pástor, Sinha, and Swaminathan (2008), Rytchkov (2008), Campbell and Thompson (2008), Lettau and Van Nieuwerburgh (2008), Ang and Bekaert (2007), Lettau and

Ludvigson (2005), Brennan and Xia (2005) and Menzly, Santos and Veronesi (2002). While these papers focus on aggregate present value models, the key to our analysis is incorporating cross-sectional information. Vuolteenaho (2002), Hansen, Heaton, and Li (2008), Pástor and Veronesi (2003, 2006), Kiku (2006), and Kelly (2011) also model valuation ratios for individual assets, though we are the first to exploit a factor structure in value ratios to form market return and cash flow forecasts.

The economics literature mainly relies on principal components (PCs) to condense information from the large cross section into a small number of predictive factors *before* estimating a linear forecast, an approach exemplified in the macro-forecasting literature by Stock and Watson (2002) and Bai and Ng (2006). PC forecasts based on macroeconomic indicators have recently been applied in the context of stock return prediction by Ludvigson and Ng (2007). The key difference between principal components and partial least squares is their method of dimension reduction. PLS condenses the cross section according to *covariance with the forecast target* and chooses a linear combination of predictors that is optimal for forecasting. On the other hand, PC condenses the cross section according to *covariance within the predictors*. The principal components that best describe predictor variation are not necessarily the factors most useful for forecasting, and therefore PCs can produce suboptimal forecasts (see Kelly and Pruitt (2011) for a detailed discussion). As we show, principal components have little forecasting success in our present value setting.

PLS is reminiscent of two-pass regression used in tests of cross-sectional beta-pricing models (see Fama and MacBeth 1973, Shanken 1992 and Jagannathan and Wang 1998).[5] Both techniques rely on cross-sectional dispersion of financial variables to infer market risk prices. There are two key differences between our three-pass regression approach and two-pass return tests. First, our cross section is constituted of valuation ratios rather than returns. Second, we string together period-by-period estimates from second-stage regressions to construct our key predictor variable, as opposed to averaging second-stage output over time to find a single

[5]Kelly and Pruitt (2011) formalize the sense in which the Fama-MacBeth procedure can be interpreted as a latent factor estimator called the three-pass regression filter.

risk price. Our approach is also related to Polk, Thompson and Vuolteenaho (2006) who use a CAPM-motivated two-pass approach to forecasting returns, and to Chowdhry, Roll, and Xia (2005), who construct estimates of the latent inflation time series from the cross section of returns using two-pass regression.

In the next section we present an economic framework for the cross section of present values. Section II introduces partial least squares and relevant results from Kelly and Pruitt (2011). In Section III we present empirical findings, compare alternative methodologies and discuss our results. We present our conclusions in Section IV. Details about PLS, as well as additional proofs, technical assumptions and other details, are relegated to the appendix.

I The Cross-Sectional Present Value System

We assume that one-period expected log returns and log cash flow growth rates across assets and over all horizons are linear in a set of common factors[6]

$$\begin{aligned} \mu_{i,t} &= \mathbb{E}_t[r_{i,t+1}] \quad = \gamma_{i,0} + \boldsymbol{\gamma}_i' \boldsymbol{F}_t \\ g_{i,t} &= \mathbb{E}_t[\Delta cf_{i,t+1}] = \delta_{i,0} + \boldsymbol{\delta}_i' \boldsymbol{F}_t + \varepsilon_{i,t}. \end{aligned} \tag{3}$$

Equation 3 states that, conditional on time t information, expected one-period returns and growth rates are driven solely by the $(K_F \times 1)$ vector of all common factors $\boldsymbol{F}_t$. For $\mu_{i,t}$ any $K_f \leq K_F$ factor loadings can be non-zero and hence the remaining $K_F - K_f$ factors do not explain expected returns; an analogous situation exists for $g_{i,t}$.

To emphasize the parsimony of our approach, we will focus on the case $K_f = 1$ (though our approach generalizes to multiple factors). The factor loading vectors $\boldsymbol{\gamma}_i$ and $\boldsymbol{\delta}_i$ sum-

[6]Factor models are analytically tractable and are sufficiently general to subsume a wide range of models considered in the asset pricing literature. Asset pricing models, both theoretical and empirical, link individual expected returns to aggregate expected returns either directly, as in the CAPM (Sharpe 1964, Lintner 1965, Treynor 1961) and Fama-French model (Fama and French 1993), or indirectly via common state variables, as in Merton's (1973) ICAPM. Similarly, theoretical models commonly assume a factor structure in dividend growth (Connor 1984, Bansal and Viswanathan 1993, and Bansal, Dittmar, and Lundblad 2005, among others).

marize how market expectations respond to movements in the underlying economic factors. We assume that assets' expected returns are determined by systematic factors and possess no idiosyncratic behavior.[7] This restriction is not imposed for individual asset's expected dividend growth, which may possess an idiosyncratic component, $\varepsilon_{i,t}$. The aggregate market obeys the same structure, with no i subscripts, and we impose that the factor model is exact for aggregate dividend growth:

$$\begin{aligned} \mu_t &= \gamma_0 + \boldsymbol{\gamma}' \boldsymbol{F}_t \\ g_t &= \delta_0 + \boldsymbol{\delta}' \boldsymbol{F}_t. \end{aligned} \tag{4}$$

Realized returns and growth rates are equal to their conditional expectations plus an unforecastable shock:

$$\begin{aligned} r_{i,t+1} &= \mu_{i,t} + \eta^r_{i,t+1} \\ \Delta cf_{i,t+1} &= g_{i,t} + \eta^c_{i,t+1}. \end{aligned}$$

Finally, we assume that the factor vector evolves as a first order vector autoregression[8]

$$\boldsymbol{F}_{t+1} = \boldsymbol{\Lambda}_1 \boldsymbol{F}_t + \boldsymbol{\xi}_{t+1}. \tag{5}$$

The above structure may be embedded in the linearized present value formula of Vuolteenaho (2002). This accounting-based identity relates an asset's log book-to-market ratio to future discount rates and earnings growth

$$v_{i,t} = \frac{\kappa_i}{1-\rho_i} + \sum_{j=1}^{\infty} \rho_i^{j-1} \mathbb{E}_t[-r_{i,t+j} + \Delta cf_{i,t+j}].$$

[7] This assumption can be relaxed. Allowing for an orthogonal idiosyncratic component in firms' expected returns has no impact on the development or implementation of our approach.

[8] That $\boldsymbol{F}_t$ is a first order process is without loss of generality since any higher order vector autoregression can be written as a VAR(1).

where $v_{i,t}$ is the log book-to-market ratio of stock i, $r_{i,t+j}$ is its log return, $\Delta cf_{i,t+j}$ is its return on equity (ROE), and κ_i and ρ_i are linearization constants. ROE is defined as[9]

$$ROE_{t+j} = \log\left(1 + \frac{\text{earnings}_{t+j}}{\text{book equity}_{t+j-1}}\right).$$

This weighted sum of expected one-period returns and growth rates over all future horizons, combined with (3) and (5), reduces to the following expression for the *ex ante* valuation ratio

$$v_{i,t} = \phi_{i,0} + \boldsymbol{\phi}_i' \boldsymbol{F}_t + \varepsilon_{i,t}, \tag{6}$$

where formulas for $\phi_{i,0}, \boldsymbol{\phi}_i$ and $\varepsilon_{i,t}$ are provided in Appendix A.A. Equations 4 and 6 unify disaggregated valuation ratios and aggregate expectations via a common factor model. They also provide a framework for utilizing cross section information to achieve our ultimate goal of precisely estimating conditional expected market returns and cash flow growth.

An alternative to Vuolteenaho's (2002) present value system is the well known Campbell and Shiller (1988) present value identity, which relates the log price-dividend ratio of an asset to its future discount rates and dividend growth. The Campbell-Shiller identity also falls into the framework of Equations 3-6 when $v_{i,t}$ is the log price-dividend ratio, $r_{i,t+j}$ is the log return, and $\Delta cf_{i,t+j}$ is log dividend growth. Fama and French (2000) show a steep downward trend the fraction of U.S. firms paying dividends, with only 20.8% of firms classified as dividend payers in 1999. Because price-dividend ratios are undefined for the majority of stocks, our analysis focuses on the cross section of book-to-market ratios. However, we do consider the performance of price-dividend ratios as a robustness check.

[9] Vuolteenaho (2002) represents this identity in terms of excess returns. We use his identity exactly, though we represent it in terms of returns rather than excess returns.

II Estimation

In this section we outline our empirical methodology, which is based on Kelly and Pruitt's (2011) generalization of partial least squares. Interested readers can refer to that paper for detailed econometric development and proofs of results stated below. Assumptions underlying the stated results are explained here and made precise in Appendix A.B.

II.A Setup

To ease the algebraic development we first establish notation. Partial least squares forecasts use two sets of inputs. The first input is the forecasting target, which in general takes the form $y_{t+1} = \beta_0 + \boldsymbol{\beta}'\boldsymbol{F}_t + \eta_{t+1}$. We will focus primarily on two targets, aggregate market returns r_m and cash flow growth Δcf_m, implying

$$y_{t+1} = \begin{cases} \gamma_0 + \boldsymbol{\gamma}'\boldsymbol{F}_t + \eta^r_{t+1} & \text{if} \quad y_{t+1} = r_{m,t+1} \\ \delta_0 + \boldsymbol{\delta}'\boldsymbol{F}_t + \eta^c_{t+1} & \text{if} \quad y_{t+1} = \Delta cf_{m,t+1}. \end{cases}$$

Defining $\underset{(T\times K_F)}{\boldsymbol{F}} = [\boldsymbol{F}_1, \boldsymbol{F}_2, ..., \boldsymbol{F}_T]'$, the matrix representation of y_{t+1} is

$$\begin{aligned} \underset{(T\times 1)}{\boldsymbol{y}} &= [y_2, y_3, ..., y_{T+1}]' \\ &= \boldsymbol{\iota}\beta_0 + \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{\eta} \end{aligned}$$

where $\beta_0, \boldsymbol{\beta}$ are defined in the obvious way for either $r_{m,t+1}$ or $\Delta cf_{m,t+1}$. The second input is the cross section of book-to-market ratios $v_{i,t} = \phi_{i,0} + \boldsymbol{\phi}_i'\boldsymbol{F}_t + \varepsilon_{i,t}$ $(i = 1, ..., N)$. These are arranged into the vector $\boldsymbol{x}_t = (v_{1,t}, ..., v_{N,t})'$ and stacked as

$$\begin{aligned} \underset{(T\times N)}{\boldsymbol{X}} &= [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T]' \\ &= \boldsymbol{\iota}\boldsymbol{\phi}_0' + \boldsymbol{F}\boldsymbol{\Phi}' + \boldsymbol{\varepsilon} \end{aligned}$$

with $\underset{(N\times 1)}{\boldsymbol{\phi}_0} = [\phi_{1,0}, \phi_{2,0}, ..., \phi_{N,0}]'$ and $\underset{(N\times K_F)}{\boldsymbol{\Phi}} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, ..., \boldsymbol{\phi}_N]'$.

II.B The Estimator

PLS can be implemented by the following series of ordinary least squares regressions. In the first stage, we run a time series regression of the book-to-market ratio for each stock i on the forecast target

$$v_{i,t} = \hat{\phi}_{i,0} + \hat{\boldsymbol{\phi}}_i' y_{t+1} + e_{i,t}.$$

The resulting estimate $\hat{\boldsymbol{\phi}}_i$ describes the sensitivity of each $v_{i,t}$ to the latent factor driving the forecast target.

In the second stage, for each period t, we run a cross-sectional regression of assets' book-to-market ratios on their loadings estimated in the first stage

$$v_{i,t} = \hat{c}_t + \hat{\boldsymbol{F}}_t{}' \hat{\boldsymbol{\phi}}_i + w_{i,t}.$$

Here, the first stage loadings become the independent variables, and the latent factors $\boldsymbol{F}_t$ are the coefficients to be estimated. The first two stages exploit the factor nature of the system to draw inferences about the underlying factors. As the factors fluctuate over time, the cross section of valuation ratios fans out or compresses. If the true factor loadings $\boldsymbol{\phi}_i$ were known, we could consistently estimate the latent factor time series by simply running cross section regressions of $v_{i,t}$ on $\boldsymbol{\phi}_i$ period-by-period. Since $\boldsymbol{\phi}_i$ is unknown, the first-stage regression coefficients provide a preliminary description of how each $v_{i,t}$ depends on $\boldsymbol{F}_t$. This first stage regression sketches a map from the cross-sectional distribution of value ratios to the latent factors. Second-stage cross section regressions of $v_{i,t}$ on first-stage coefficients use this map to produce estimates of the factors at each point in time.

The third step in the filter runs a predictive regression of realized returns or cash flow growth rates on the lagged factors estimated in the second stage. This final regression is the culmination of the multi-asset present value system. It parsimoniously combines information

from individual assets' valuation ratios to arrive at a prediction of future aggregate returns and dividend growth. The ultimate predictor, $\hat{\boldsymbol{F}}_t$, is a discerningly-constructed linear combination of disaggregated price-dividend ratios that collapses the cross section system to its fundamental driving factors. The R^2 from the final step regression summarizes the predictive power of the multi-asset present value model; that is, the predictive accuracy of market expectations embodied in valuation ratios.

Kelly and Pruitt (2011) provide a one-step representation of this algorithm:

$$\hat{\boldsymbol{y}} = \boldsymbol{\iota}\bar{y} + \boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{y}\left(\boldsymbol{y}'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{y}\right)^{-1}\boldsymbol{y}'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{y}, \quad (7)$$

where $\boldsymbol{J}_L \equiv \boldsymbol{I}_L - L^{-1}\boldsymbol{\iota}_L\boldsymbol{\iota}_L'$, $\boldsymbol{I}_L$ is the L-dimensional identity matrix and $\boldsymbol{\iota}_L$ is a L-vector of ones. $\boldsymbol{J}$ matrices are present since each regression step is run including a constant regressor.

This procedure is consistent: It asymptotically recovers the latent expectations of aggregate market returns and cash flow growth as the number of predictors and time series observation both become large.[10] In particular, Theorems 1 and 4 in Kelly and Pruitt (2011) show that $\hat{\beta}_0 + \boldsymbol{\beta}'\hat{\boldsymbol{F}}_t$ is normally distributed around $\mathbb{E}_t[y_{t+1}]$ under the assumptions in Appendix A.B as $N, T \to \infty$. These assumptions are quite weak. The key assumption is that log book-to-market ratios obey a linear factor structure, which is consistent with a range of theories for conditional expected returns (assuming that ROE is also linear in its factors). The remaining assumptions are largely technical, and impose that second moments are finite and probability limits are well-behaved, that there is limited time series and cross-sectional autocorrelation among elements of the residual matrices $\boldsymbol{\eta}$ and $\boldsymbol{\varepsilon}$, and impose that innovations to returns and dividend growth are asymptotically orthogonal to lagged factors and value ratios.

The general version of our theory accommodates multiple factors in both returns and

[10] Because the first-stage regression takes an errors-in-variables form, coefficients estimated in the first and second stage are biased. The third regression step accounts for this bias, and removes this effect from the ultimate forecast of y, since least squares fitted values are invariant to scalar multiples of regressors or additive constants. See the consistency argument in Kelly and Pruitt (2011) for details.

cash flow growth. In the interest of parsimony, and to highlight the power of our approach compared to the large set of alternative univariate predictors, we assume that returns and cash flow expectations are each driven by a single factor (though these factors are allowed to be different for returns versus growth rates). Extending our analysis to extract additional factors from the cross section of valuation ratios transforms our third step forecasts into multivariate predictive regressions, rather than univariate predictions, and can potentially improve forecastability beyond what we document below.

II.C In-Sample Versus Out-of-Sample Implementation

Throughout our empirical analysis we consider both in-sample and out-of-sample approaches to implementing our forecasts. To outline the differences in information sets it is convenient to work within the three-stage regression construction of the filter rather than the direct formulation in (7).

The basic implementation, which uses all available information, is a purely in-sample estimation. First-stage regressions use the full time series of data to estimate factor loadings. Second-stage regressions that produce the predictor variable at each time t use only price-dividend ratio data at time t and constant factor loadings estimated in the first stage. Our predictive factor is exactly this second-stage regression and therefore is a linear combination of time t data that does not contain look-ahead bias. Third-stage predictive regressions are also run in-sample.

In the full information version it is possible that first-stage regressions introduce a small sample bias in our predictors since first-stage factor loadings are based on the full time series. This is analogous to small sample bias in standard OLS predictive regression (cf. Stambaugh (1986) and Nelson and Kim (1993)), which enters into forecasts via estimated predictive coefficients.[11] Consider, for instance, OLS forecasts of r_{t+1} on some predictor z_t, where both r and z are mean zero. The in-sample estimated coefficient is $\hat{b}_T = (\frac{1}{T}\sum_{t=0}^{T-1} r_{t+1}z_t)/(\frac{1}{T}\sum_{t=0}^{T-1} z_t^2)$.

[11] Small sample bias also arises from the preliminary parameter estimation step of a Kalman-filtered state space.

The forecast for r_{t+1} is given by $\hat{b}_T z_t$, thus the targeted observation is directly contributing through the parameter estimate an order of $1/T$ of the total information in its own forecast. For small T, this can favor false detection of predictability. However, this bias is truly a small sample phenomenon because each future observation has vanishing importance for first-stage coefficient estimates. Our annual forecasts consists of overlapping observations spanning up to 80 years, while our monthly forecasts provide up to 960 non-overlapping time series observations. While neither of these sample sizes is particularly small, we nonetheless take the possibility of small sample bias very seriously.

Thus, our second implementation is a pure (recursive) out-of-sample analysis, and these results serve as the focus of our empirical analysis. The procedure we use is a standard out-of-sample estimation scheme which has been well-studied in the literature (e.g. Goyal and Welch (2008)). The main idea is to run first- and third-stage estimations on training samples that exclude the return or cash flow observation ultimately forecast. We split the full T-period sample at date τ, using the first τ observations as a *training* sample and the last $T - \tau$ observations as the *left-out* sample. We estimate first-stage factor loadings using observations $\{1, ..., \tau\}$. Then, for each period $t \in \{1, ..., \tau\}$, we estimate the time t value of our predictor variable using the cross section of valuation ratios at t and first-stage coefficients (which are based on data $\{1, ..., \tau\}$). We then estimate the predictive coefficient in a third-stage forecasting regression of realized returns (or cash flow growth) for periods $\{2, ..., \tau\}$ on the factor extracted from $\{1, ..., \tau - 1\}$. Finally, our out-of-sample forecast of the $\tau + 1$ return is the product of the third-stage predictive coefficient and the time τ second-stage result. This procedure is then repeated recursively, next using only data from $\{1, ..., \tau + 1\}$, to construct a forecast for the return at $\tau + 2$, until the entire sample of length T has been exhausted.

All of our analyses are performed using data at the monthly frequency. The out-of-sample procedure just described may be applied to one month forecasts without modification. Annual out-of-sample forecasts require additional care to account for overlap in monthly

observations. Hence for annual returns we use the training sample $\{1, ..., \tau\}$ to estimate all model coefficients for the purpose of forecasting the annual return though $\tau + 12$. The resulting annual horizon forecasts are genuinely out-of-sample.

Inference for in-sample one month forecasts is based on the asymptotic distributions for PLS estimates derived in Kelly and Pruitt (2011). For one year forecasts, we use overlapping data and must adjust our standard errors to reflect the dependence that this introduces into forecast error. We do this in three ways. First, we calculate Kelly and Pruitt (2011) standard errors only using forecast errors from December of each year. This avoids the overlapping observations problem and, with 80 non-overlapping observations, still provides a reasonable sample size for approximating the theoretical asymptotic test statistic distribution.[12] As a second alternative, we calculate Hodrick (1992) standard errors using all overlapping observations. This approach explicitly accounts for the moving average structure that overlap introduces into residuals. Third, we report Newey-West (1985) standard errors with twelve lags to account for overlap-induced serial correlation among residuals. In our empirical analysis, results for all of these test statistics are very similar. Our exposition focuses on the Kelly and Pruitt in-sample standard errors since our results show that these are typically the most conservative.

PLS in-sample tests have no direct out-of-sample counterpart (this is also true of OLS). Instead, we conduct out-of-sample inference with the "encompassing" forecast test ENC-NEW derived by Clark and McCracken (2001). This statistic has become widely used in the forecasting literature, and tests the null hypothesis that two predictors provide the same out-of-sample forecasting performance. When we report this statistic, we are testing the denoted predictor versus the historical mean of the target series.[13] We report significance levels as

[12] We can similarly calculate Kelly and Pruitt standard errors using non-overlapping forecast errors corresponding to any month January through December. We report the p-value for the median of 12 t-statistics constructed for each of the possible year-end months, therefore covering every non-overlapping set of forecast errors. In practice any choice of year-end month yields very similar statistics so that the median is representative.

[13] This test compares our time τ forecast of the $\tau + 1$ ($\tau + 12$ for annual returns) realization against the forecast based on the target variable's mean estimated through time τ.

found from Clark and McCracken's (2001) appendix tables, where critical values for the 10%, 5% and 1% levels are provided. The notation "$< x$" represents the smallest significance level x for which the encompassing test statistic exceeds the critical value. When evaluating overlapping forecast errors (as we do for annual return, dividend growth and earnings growth forecasts) we use Newey-West standard errors with twelve lags to consistently estimate the appropriate asymptotic variance in the denominator of ENC-NEW, as suggested in Clark and McCracken (2005). Out-of-sample results reported in tables are based on a 1980 sample split, save for the international data which are split at 1995 owing to its much more recent start date. As fore mentioned, we plot our-of-sample R^2 over a wide range of alternative split dates to demonstrate the robustness of our results to choice of sample split.

To evaluate forecasting fit, we calculate the predictive $R^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2}{\sum_t (y_t - \bar{y})^2}$, which for our PLS forecasts is equal to the R^2 of the third stage univariate regression. The out-of-sample R^2 lies in the range $(-\infty, 1]$, where a negative number means that a predictor provides a less accurate forecast than the target's historical mean.

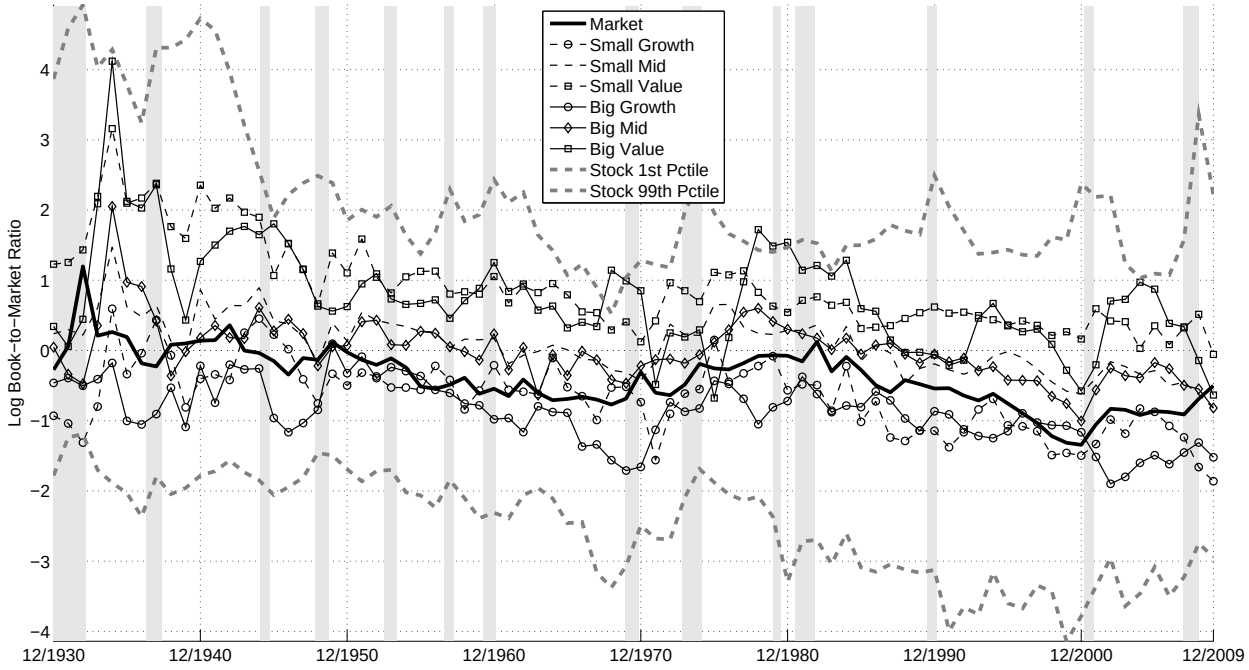
III Empirical Results

III.A Data

Our empirical analysis examines market return and cash flow growth predictability by applying partial least squares to different cross sections of valuation ratios. The main three are book-to-market ratios for Fama and French's (1993) size and value-sorted portfolios (in which U.S. stocks are divided into six, 25 or 100 portfolios).

Many authors, including Fama and MacBeth (1973), Miller and Scholes (1982), Fama and French (1988) and Polk, Thompson and Vuolteenaho (2006), have highlighted the difficulties in working noisy firm-level accounting variables (book values and cash flows). Fama and MacBeth prescribe the use portfolios to reduce the impact of individual stock noise on our information extraction, and this drives our focus on the cross section of portfolios rather

Figure 1: Book-to-Market Ratios



Notes: The figure shows year-end log book-to-market ratios from 1930-2009 for the aggregate U.S. stock market, six size and value-sorted portfolios from Fama and French (1993) (available on Kenneth French's website), and the first and 99^{th} percentile (each year) for individual U.S. stock book-to-market ratios (data from CRSP and Compustat). NBER recession dates are represented by the shaded area.

than individuals stocks for our main analysis.[14] However, we also analyze the robustness of our main results to several other cross sections. One alternative we explore is indeed the cross section of individual firm data, at which point we also discuss in more detail the tradeoffs associated with portfolio data versus that of individual firms. We also consider price-dividend ratios for size and value-sorted portfolios in place of book-to-market ratios. Finally, we take our analysis to international data, using the country-level portfolio valuation ratios of Fama and French (1998). Our focus is on the 1930-2009 sample for U.S. data. The international sample is available only from 1975-2009. Individual stock data are from CRSP and Compustat. U.S. and international portfolio data are from Kenneth French's website. Alternative predictors are obtained from data on Amit Goyal's website.

[14]Our model implies that it is the similarity of the factor loadings ($\boldsymbol{\gamma}$ or $\boldsymbol{\delta}$) across firms in the same portfolio that allow for the noise reduction with respect to forecasting market aggregates.

Figure 1 plots book-to-market ratios of the aggregate U.S. stock market and six Fama-French size and value-sorted portfolios in December of each year. It also shows the 1^{st} and 99^{th} percentile of individual U.S. stock book-to-market ratios by year. Disaggregated book-to-market ratios exhibit more variability than that of the market portfolio, with a standard deviation of 0.53 on average across the six Fama-French portfolios, versus 0.42 for the market. The portfolio series as well as the individual stock percentiles show that cross sectional dispersion among book-to-market ratios varies dramatically over time. The interquartile range of stock-level log book-to-market ratios reaches its maximum of 1.8 during the Great Depression, falls to 0.7 shortly after World War II, and rises again to 1.5 during the technology boom of the late 1990s. Variation in ordering, dispersion, and comovement of book-to-market ratios is the key predictive information that PLS exploits to measure market expectations.

III.B Market Return Predictability

III.B.1 Forecasting with Portfolio Book-to-Market Ratios

Our main empirical analysis evaluates the predictability of aggregate market returns using the cross section of book-to-market ratios. We directly estimate our model of the cross section system described in Section I. This model emphasizes the low dimension predictive structure underlying the many-predictor cross section, which motivates our estimation of the model via PLS. Table I presents return forecasting results based on six, 25 and 100 book-to-market ratios of size and value-sorted portfolios of U.S. stocks (Fama and French (1993)). We consider two different forecasting horizons – one month and one year – and report findings for both in-sample and out-of-sample forecasts.

The left half of Table I shows that a single factor extracted via PLS demonstrates a striking degree of predictability for one year returns. The in-sample implementation generates a predictive R^2 reaching 25.0% and 32.9% based on 25 and 100 book-to-market ratios, respectively ($p < 0.001$ in both cases). When only six portfolios are used, the in-sample

Table I: Market Return Predictions (1930-2009)

	One Year Forecasts				One Month Forecasts	
	R^2 (%)	p (KP/CM)	p (Hodrick)	p (NW)	R^2 (%)	p (KP/CM)
			6 Portfolios			
In-Sample	3.84	0.126	0.119	0.062	0.51	0.136
Out-of-Sample	5.85	< 0.050	-	-	0.44	< 0.100
			25 Portfolios			
In-Sample	25.04	< 0.001	< 0.001	< 0.001	2.07	< 0.001
Out-of-Sample	9.79	< 0.010	-	-	0.60	< 0.010
			100 Portfolios			
In-Sample	32.87	< 0.001	< 0.001	< 0.001	4.08	< 0.001
Out-of-Sample	9.88	< 0.010	-	-	0.76	< 0.010

Notes: Results of PLS forecasts of one year and one month market returns. The sets of predictor variables are six, 25 and 100 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993). In-sample results are for the 1930-2009 sample. Out-of-sample forecasts split the sample in 1980, using the pre-1980 period as a training window, and recursively forecasting returns beginning in January 1980 (results from a wide range of alternative sample splits are shown in Figure 3). We report in-sample and out-of-sample forecasting regression R^2 in percent. We also report p-values of three different in-sample test statistics. The first is the asymptotic predictive loading t-statistic from Kelly and Pruitt (2011), calculated on every non-overlapping set of residuals as described in the text. For annual returns we also report Hodrick (1992) and Newey-West (1985) t-statistic p-values. For out-of-sample tests we report Clark and McCracken's (2001) ENC-NEW encompassing test statistic in *italics* with p-values from Clark and McCracken's (2001) appendix tables. This tests the null hypothesis of no forecast improvement over the historical mean; for annual returns we follow Clark and McCracken (2005) and use Newey-West standard errors with twelve lags to consistently estimate the appropriate asymptotic variance.

forecasting relation is only marginally significant ($R^2 = 3.84\%$, $p = 0.126$), highlighting the information gains from using finer portfolio divisions. Out-of-sample PLS forecasts based on portfolio book-to-market ratios are similarly powerful, delivering an R^2 of 5.9%, 9.8% and 9.9% for six, 25 and 100 portfolios. These are large economic magnitudes for out-of-sample prediction, comparable to in-sample results from commonly studied predictors such as the aggregate price-dividend ratio. Each of these out-of-sample results are statistically significant at least at the 5% level based on Clark-McCracken tests, and at the 1% level for finer portfolio divisions.

The last two columns report forecasting results for one month returns. The monthly in-

sample R^2 reaches 2.1% and 4.1% based on a single linear combination of 25 or 100 portfolio book-to-market ratios, respectively. Out-of-sample one month return forecasts are significant at the 1% level or better for 25 and 100 portfolios, with an R^2 reaching as high as 0.8%. At the monthly frequency, an out-of-sample R^2 of 0.8% has large economic significance. A heuristic calculation suggested by Cochrane (1999) shows that the Sharpe ratio (s^*) earned by an active investor exploiting predictive information (summarized by the regression R^2) and the Sharpe ratio (s_o) earned by a buy-and-hold investor are related by $s^* = \sqrt{\frac{s_o^2+R^2}{1-R^2}}$. Campbell and Thompson (2008) estimate a monthly equity buy-and-hold Sharpe ratio of 0.108 using data back to 1871. Therefore, an out-of-sample predictive R^2 of 0.8% implies that an active investor exploiting our approach could achieve a Sharpe ratio improvement of roughly 30% over a buy-and-hold investor, using only real-time information in the form of portfolio book-to-market ratios.

How do our market return forecasts compare with predictors proposed in earlier literature? Table II compares predictive accuracy of our approach with an extensive collection of alternative predictors that have been considered in the literature. In particular, we explore forecasts from 16 predictors studied in a recent return predictability survey by Goyal and Welch (2008). The table considers both in-sample and out-of-sample forecasts of market returns over horizons of one year and one month from each regressor individually. Among the alternatives, the best univariate forecasts at the annual horizon (Panel A) are achieved by *cay* (Lettau and Ludvigson 2001), which delivers an in-sample R^2 of 14.9% and an out-of-sample R^2 of 2.5%, showing statistically significant predictive power at the 1% level.[15] Other successful out-of-sample predictors include the cross section premium (*csp*) of Polk, Thompson and Vuolteenaho (2006), the term spread (*tms*), the long term government bond return (*ltr*), and the aggregate log earnings-to-price ratio (*ep*). The first row of the column shows that even the success of *cay* is dominated by the single PLS factor extracted from

[15]Specifically, this is the cayp variable which Goyal and Welch (2008) discuss at some length, analyzing how the typical "out-of-sample" construction of the variable actually uses information from the full sample. The highest frequency at which *cay* is available is quarterly, therefore we take each observation to represent the quarter's last month observation and treat the other months as missing.

portfolio-level book-to-market ratios.

Table II also reports forecasting results using the first three principal components extracted from the cross section of 25 portfolio book-to-market ratios.[16] Principal components fail to demonstrate any significant return forecasting power in-sample or out-of-sample. In the finance and economics literature, principal components (PC) has become the *de facto* method of for condensing large cross sections of predictors into a small number of predictive factors. PC suffers from an important shortcoming when it comes to forecasting: The components that best describe variation among the predictors are not necessarily the factors most useful for predicting next period's aggregate return. PLS, on the other hand, extracts predictive factors that are optimal for forecasting. If there is a common factor among the predictors that is useful for forecasting returns, but that describes a relatively small amount of the variation within the predictors (that is, it is a low ranking principal component), PC can fail to detect that factor. PLS differs in that it only identifies *forecast-relevant* factors, ignoring factors that may be pervasive among predictors but useless for forecasting. For an in depth discussion of PLS versus PC forecast properties, we refer readers to Kelly and Pruitt (2011).

Predictions of one month returns (Table II, Panel B) tell the same story as annual forecasts. Our procedure is the dominant in-sample univariate predictor ($R^2 = 4.1\%$, $p < 0.001$), with only *cay* and the log earnings-price ratio (*ep*) as the other predictors with in-sample forecasting power that is significant at least at the 5% level. Out-of-sample, only our procedure ($R^2 = 0.8\%$, $p < 0.01$) and the default rate (*dfr*) provide significant positive results, but the latter's in-sample predictive power is tiny and insignificant ($p = 0.5$). In summary, our PLS factor derived from the cross section of book-to-market ratios is the only predictor to exhibit significant performance both in-sample and out-of-sample for one month returns.

Estimates of the expected annual return based on our cross-sectional approach are plotted in Figure 2, and are compared against estimates from regressions on aggregate valuation

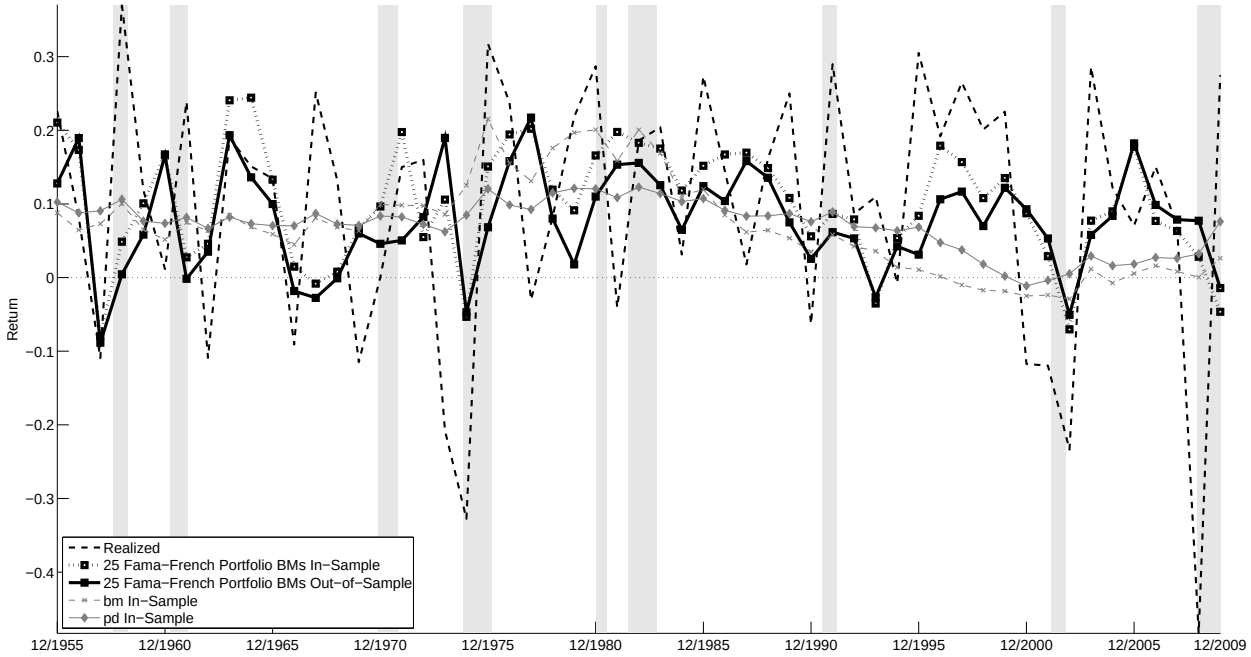
[16] Principal components results are very similar when extracted from 100 portfolios.

Table II: Market Return Predictions: Comparison with Common Alternative Predictors

	Panel A: One Year Forecasts					Panel B: One Month Forecasts			
	In-Sample			Out-of-Sample		In-Sample		Out-of-Sample	
	R^2 (%)	p (Hodrick)	p (NW)	R^2 (%)	p (CM)	R^2 (%)	p (NW)	R^2 (%)	p (CM)
100 Ptfs	**32.87**	**< 0.001**	**< 0.001**	**9.88**	**< 0.010**	**4.08**	**< 0.001**	**0.76**	**< 0.010**
dfy	0.31	0.674	0.664	-0.13	-	0.05	0.778	-0.39	-
infl	0.00	0.853	0.747	-0.31	-	0.04	0.641	-0.14	-
svar	0.00	0.995	0.650	-0.13	-	0.09	0.654	-0.80	-
csp	0.36	0.401	0.492	1.29	< 0.050	0.61	0.043	-0.11	-
de	0.80	0.527	0.525	-5.59	-	0.09	0.574	-1.54	-
lty	0.79	0.351	0.279	-0.73	-	0.02	0.671	-0.87	-
tms	1.10	0.195	0.151	0.41	< 0.100	0.12	0.266	-0.20	-
tbl	0.16	0.680	0.661	-3.74	-	0.00	0.962	-1.03	-
dfr	0.00	0.553	0.614	-0.27	-	0.15	0.499	1.06	< 0.010
pd	3.57	0.090	0.062	-3.08	-	0.20	0.304	-7.44	-
dy	3.84	0.080	0.047	-6.31	-	0.32	0.200	-0.23	-
ltr	0.89	0.002	< 0.001	0.89	< 0.050	0.37	0.067	-5.16	-
ep	8.04	< 0.001	< 0.001	1.35	< 0.010	0.47	0.044	-3.48	-
bm	9.99	0.002	< 0.001	-18.67	-	0.81	0.096	-3.58	-
ntis	8.19	0.081	0.071	-56.44	-	0.52	0.110	-2.13	-
cay	14.93	0.004	< 0.001	2.48	< 0.010	1.59	0.009	-2.32	-
pc1	1.96	0.182	0.115	-2.19	-	0.16	0.204	-0.62	-
pc2	0.08	0.811	0.790	-0.32	-	0.00	0.857	-0.03	-
pc3	0.02	0.960	0.882	-0.55	-	0.03	0.764	0.00	-

Notes: Results of PLS forecasts of one year and one month market returns from 100 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993); and results for alternative predictors taken from Goyal and Welch (2008) with data updated through 2009. These alternative predictors are the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp), the dividend payout ratio (de), the long term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the price-dividend ratio (pd), the dividend yield (dy), the long term rate of returns (ltr), the earning price ratio (ep), the book to market ratio (bm), the investment to capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the *ex post* consumption-wealth-income ratio (cay) which Goyal and Welch refer to as "cayp." Additionally we consider the first three principal components extracted from the 25 portfolio book-to-market ratios and find no significant improvement when all three predictors are used together. In-sample results are for the 1930-2009 sample wherever possible. Out-of-sample forecasts split the sample in 1980, using the pre-1980 period as a training window, and recursively forecasting returns beginning in January 1980 (results from a wide range of alternative sample splits are shown in Figure 3). We report in-sample and out-of-sample forecasting regression R^2 in percent. We also report p-values of two different in-sample test statistics: Hodrick (1992) and Newey-West (1985) t-statistic p-values. For monthly returns, the Newey-West standard error is identical to the White robust standard error. For out-of-sample tests we report Clark and McCracken's (2001) ENC-NEW encompassing test statistic in *italics* with p-values from Clark and McCracken's (2001) appendix tables. This tests the null hypothesis of no forecast improvement over the historical mean; for annual returns we use Newey-West standard errors with twelve lags to consistently estimate the appropriate asymptotic variance in the denominator of ENC-NEW, as suggested in Clark and McCracken (2005).

Figure 2: Market Return Predictions



Notes: The figure shows year-end realized returns for the aggregate U.S. stock market (Realized), both in-sample and out-of-sample forecasts from our PLS factor extracted from 25 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993) (25 Fama-French Portfolio BMs In-Sample/Out-of-Sample), and in-sample forecasts from predictive regressions on the aggregate book-to-market (bm In-Sample) and aggregate price-dividend (pd In-Sample) ratios. NBER recession dates are represented by the shaded area.

ratios.[17] Our estimated one-year-ahead expected return process differs from other estimates in qualitatively important ways. Table III presents the volatility and persistence of expected market returns estimated from predictive regressions based on the aggregate market book-to-market ratio, compared to estimates based on a single PLS factor extracted from the cross section of 25 portfolio book-to-market ratios. Our estimates suggest that expected returns are nearly twice as volatile as previously estimated. We find expected return volatility of 10.8% at the annual frequency based on in-sample estimates, suggesting that about half of the annual variation in stock market value during this period is attributable to fluctuations in investor expectations. Examining the 1955-2009 subsample allows us to compare with out-of-sample estimates from our approach. During this period we see that in-sample and

[17]Though we have these predictions monthly, we plot them at an annual frequency to make the figure easy to read.

Table III: Volatility and Autocorrelation of Expected Market Return Estimates

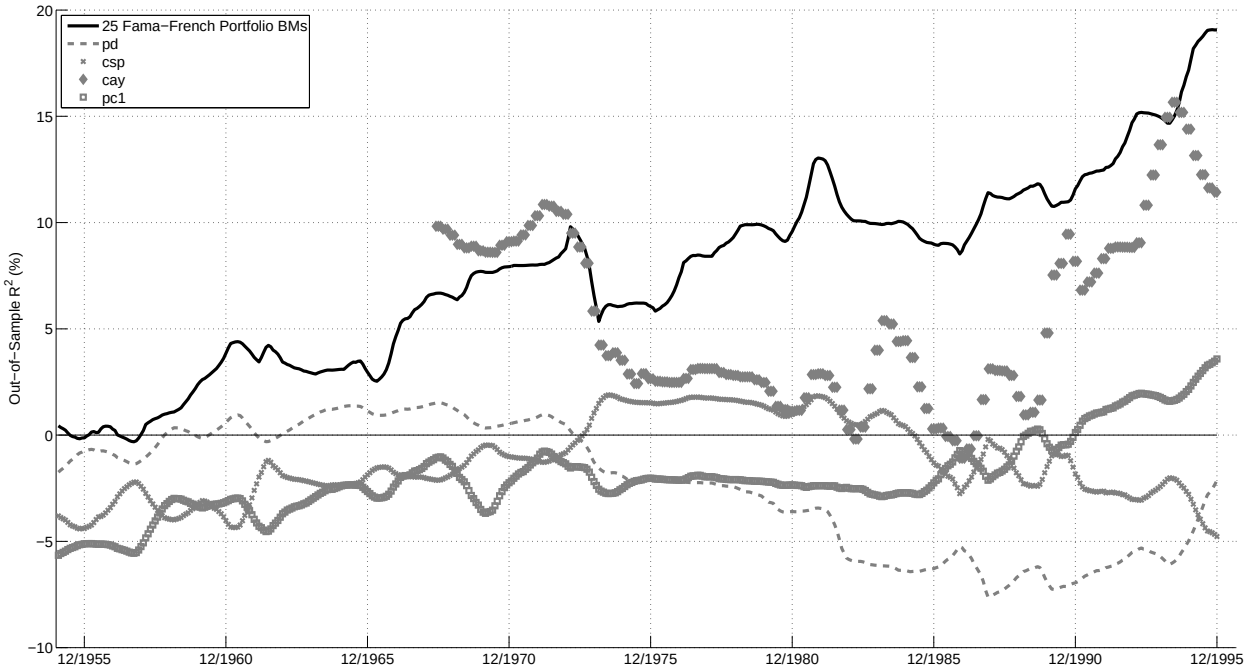
	1930-2009		1955-2009	
	Vol (%)	AC(1)	Vol (%)	AC(1)
Realized Returns	20.2	0.032	17.5	-0.113
Aggregate Price-Dividend	4.0	0.889	5.1	0.908
Aggregate Book-to-Market	6.5	0.839	3.9	0.912
25 Portfolios (in-sample)	10.8	0.332	6.7	0.445
25 Portfolios (out-of-sample)	-	-	6.9	0.199

Notes: Volatility (standard deviation) and first-order autocorrelations for realized aggregate U.S. stock market annual returns, and expected annual returns estimated using the aggregate price-dividend ratio in-sample, the aggregate book-to-market ratio in-sample or our PLS factor extracted either in-sample or out-of-sample from 25 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993). Out-of-sample expected returns begin only in 1955.

out-of-sample expected return estimates have similar volatility, nearly 7% per annum, which is about 80% higher than in-sample estimates based on aggregate value ratio and accounts for about 40% of annual stock market variation.

Our estimates also show that one-year-ahead expected returns mean revert more quickly than previously believed. We find an autocorrelation between 0.199 and 0.445 at the annual frequency, compared to aggregate value ratios that imply a persistence of 0.912. In light of the volatility and lack of serial correlation in realized returns, higher volatility and lower persistence of our one-year-ahead expected returns contribute to return forecasts that are substantially more accurate than alternative forecasts. The close agreement across in-sample and out-of-sample estimates imply that this conclusion is a genuine feature of market prices rather than an artifact of statistical overfit. Our results point to market expectation dynamics that are quite different than the persistent expected returns generated by standard consumption-based asset pricing paradigms such as habit persistence or long run risks, models that are typically calibrated to match the much weaker empirical return predictability generated by the aggregate price-dividend ratio.

Figure 3: Out-of-Sample R^2 by Sample Split Date, One Year Returns



Notes: Out-of-sample R^2 by varying sample split date. For forecasts of one year market returns from: a single PLS factor from 25 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993); the aggregate price-dividend ratio; and alternative predictors taken from Goyal and Welch (2008) – the cross-section premium (csp) and the *ex post* consumption-wealth-income ratio (cay); and the first principal component extracted from the 25 Fama-French portfolio book-to-market ratios (pc1)

III.B.2 Varying Out-of-Sample Sample Splits

Our first robustness check recognizes the following: We have reported out-of-sample forecasting tests based on a 1980 sample-split date, but recent forecast literature suggests that sample splits themselves can be data-mined (cf. Hansen and Timmermann (2011) and Inoue and Rossi (2011)). To demonstrate the robustness of out-of-sample forecasts to alternative sample splits, Figure 3 plots out-of-sample annual return predictive R^2 as a function of sample-split date for a variety of predictors. The earliest sample split we consider is January 1955, which uses less than one third of the data (25 out of 80 years) as a training sample. The latest split we consider is 1995, which uses a long training sample and limited test sample. The figure shows our procedure consistently outperforms alternative predictors across sample splits. Note that the aggregate book-to-market ratio is not shown since its

out-of-sample R^2 consistently falls below -10%. Forecasts using *cay* are competitive in only a small subset of the sample splits. The first principal component ("pc1" in the figure) has consistently poor out-of-sample performance, as do forecasts that use the first three principal components simultaneously (not shown due to close similarity with results from the first principal component). The remaining predictors fail to consistently demonstrate out-of-sample predictability across various split dates. An attractive feature of our estimator is that its out-of-sample R^2 shows a gradual, steady increase as the training sample expands. This suggests that PLS successfully learns from new information being fed into the procedure as more data becomes available, allowing it to more effectively counteract sample noise.

III.B.3 Subsample Cross-Validation

We next evaluate the robustness of our results to the alternative out-of-sample forecast procedure of cross-validation. The attractiveness of our benchmark recursive out-of-sample procedure is that it strictly relies only on information available to analysts in real time. One limitation of this test is that it always trains on an early portion of the sample and tests on a later portion. If there are differences in predictability between early and late in the sample, this can be missed by the recursive, forward looking approach. Cross-validation breaks this strict timing: Training samples need not contain the earliest observations and the test sample need not be restricted to latter portions of the data.[18] Hence cross-validation is an out-of-sample procedure that does not require temporal ordering of training and test samples. Therefore, it allows us to measure the out-of-sample performance of our forecasts for every period of our entire sample, while still requiring that model parameter estimates are *not* based on data in the left-out sample.

We begin by partitioning the data into T subsamples of length K months. For the subsample indexed by date t, we perform the full three-pass PLS procedure using all data *except* that for periods t through $t + K$. Parameter estimates are based on the training sample

[18]Stock and Watson (2011) advocate cross-validation in the context of macroeconomic forecasting by a similar rationale and use a similar scheme.

Table IV: Cross-Validation Out-of-Sample R^2

		Number of Years Left-Out				
		1	2	3	5	10
6 Portfolios	CV R^2	-4.71	-2.91	-2.43	-3.48	-2.99
25 Portfolios	CV R^2	4.85	6.62	7.19	4.63	2.33
100 Portfolios	CV R^2	3.74	4.72	5.67	6.67	11.14

Notes: R^2 for annual returns from an alternative out-of-sample procedure: Leave-K-out cross-validation, where K is twelve times the number of years dropped. Results from PLS forecasts of one year market returns from six, 25 or 100 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993). The procedure is described in more detail in the text. It differs from our main (recursive) out-of-sample results in that the training sample can include observations both temporally before and after the left-out sample.

$\{1, ..., t-1, t+K+1, ..., T\}$, including weights in the PLS factor construction and the final stage predictive coefficient.[19] These parameters estimates are then used to construct the PLS factor and forecast the annual return realized in period $t+12$.[20] This is repeated for each subsample, and hence out-of-sample forecast errors for all $t = 1, ..., T$ are aggregated to calculate a cross-validation R^2, denoted CV R^2 and reported in Table IV.[21] For all subsample window lengths, our single PLS factor forecasts from 25 and 100 portfolios produces a positive cross-validation R^2.[22] This sensitivity analysis suggests that our main out-of-sample predictability results are representative of the behavior of value ratios and market returns across various subsamples of the entire sample period.

[19] There are slight but obvious modifications required near the endpoints as t gets within K periods of the beginning or end of our sample. For example the forecast for observation 1 has $\{t+K+1, ..., T\}$ as its training sample, as clearly there is no data available prior to observation 1.

[20] That the first forecast occurs at $t+12$ reflects care taken to ensure that no information from the testing window is used in parameter estimation. This accounts for overlapping monthly observations of annual returns, as discussed in Section II.C.

[21] This procedure produces a unique forecast error for each data point. We use the calculation CV $R^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2}{\sum_t (y_t - \bar{y})^2}$, where $\bar{y}$ is mean return in the training sample corresponding to forecast $\hat{y}_t$. As in the recursive case, this R^2 has the interpretation of measuring out-of-sample predictive power relative to that of the training sample mean return.

[22] Statistical tests for cross-validation are not well-known in the economics forecasting literature. A natural alternative for such a test is the Clark and McCracken's (2001) statistic developed for recursive forecasting. Treating this as a cross-validation test, we find that out-of-sample forecasts from 25 and 100 portfolios are statistically significant at the 5% level or better in all cases.

III.B.4 Forecasting From Individual Stock Valuation Ratios

We next investigate the usefulness of information in individual stock valuation ratios for predicting market returns. We tackle the formidable task of applying our simple PLS forecasting approach directly to stock-level valuation ratios of all CRSP stocks from 1930-2009.

Polk, Thompson and Vuolteenaho (2006) is the benchmark study for combining individual firm data into a market return prediction. Their paper emphasizes the difficulty in working with noisy firm-level data, such as book value or cash flows, which may distort valuation ratios and complicate extraction of market expectations from the cross section. They address this with a series of pre-filtering adjustments to firms' valuation ratios and robust statistics. These include relying on ordinal ranks rather than cardinal values, value-weighting observations, and censoring extreme observations.

One alternative solution to this problem, as suggested by Fama and MacBeth (1973) and done in our main analysis above, is to combine individual stocks into portfolios. In this section, we are interested in drilling deeper to understand how much we can learn about aggregate market expectations directly from individual stock data.

To this end, we take another approach to dealing with noise in individual stock valuation ratios that was originally proposed in Miller and Scholes (1982) and Fama and French (1988). Those authors suggest that, rather than using infrequent and potentially mis-measured balance sheet data, it may be beneficial to omit fundamentals information entirely and focus only on the price portion of the valuation ratio. We conduct our main individual stock analysis taking this approach. To achieve stationarity in the ratio, we divide the firm's month-end share price by the moving average price over the previous three years. As in our book-to-market calculation, we then take the log of this ratio as our definition of $v_{i,t}$ in system (6). We call this the individual stock price-to-moving-average (PMA) ratio. Individual stock data comprises a severely unbalanced panel. To address this, as well as potential parameter instability at the stock level, we estimate the model using a rolling five-year estimation window, and only include stocks in each period's forecast if they have no missing observations

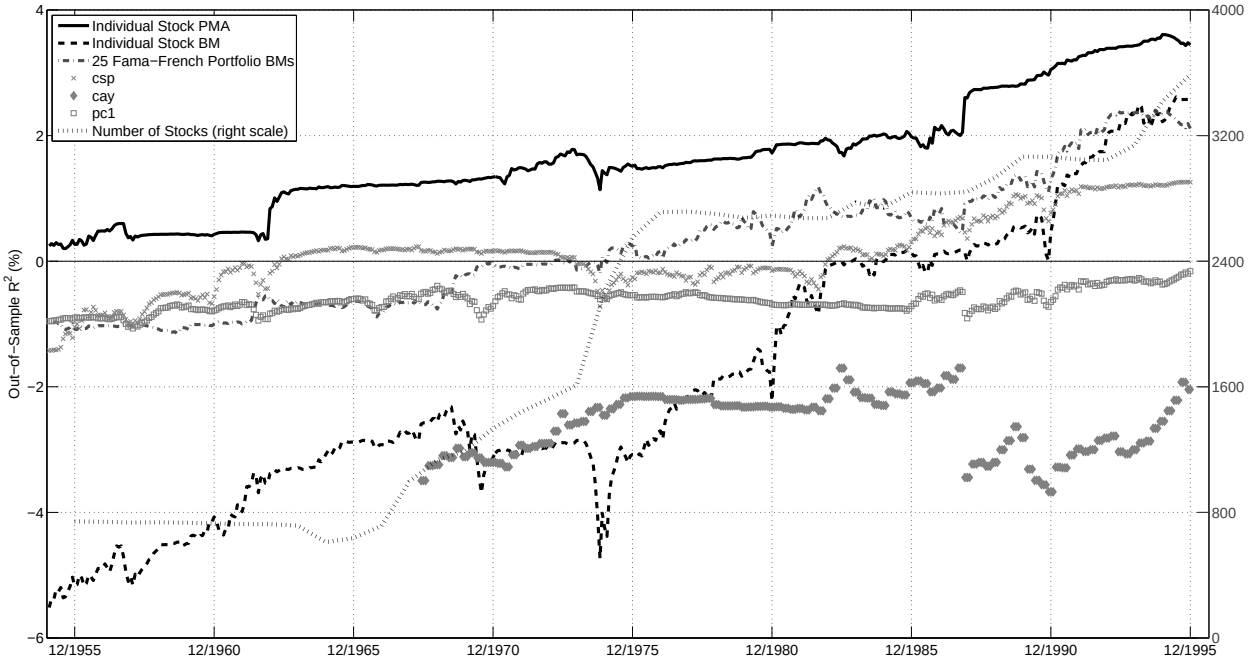
in this window.

We focus on monthly returns to directly compare with the benchmark of Polk et al., who study monthly forecasts. Figure 4 reports the out-of-sample monthly return forecasting R^2 from individual stock PMA ratios across a range of sample splits. It also plots the R^2 for Polk et al.'s *csp* variable. Our single PLS factor extracted from the cross section of individual stock PMA ratios consistently produces a positive out-of-sample R^2, rising above 1% in the mid-1960s and exceeding 2% per month by the mid-1980s. It uniformly dominates *csp*, as well as forecasts from *cay* and the first component of 25 Fama-French portfolios' book-to-market ratios.[23] For comparison, we also plot results from the PLS factor extracted from 25 portfolio book-to-market ratios.

The results of the same forecasting analysis using a single PLS factor extracted from individual stock book-to-market ratios are also plotted in Figure 4. For small training samples these forecasts produce a negative R^2, presumably due to the noisiness of firm-level balance sheet data. In early sample splits the book-to-market ratios are not only dominated by forecasts based solely on firm-level prices, but also by *csp*, whose clever modifications mitigate the influence of noise in firm-level valuation ratios using balance sheet data. Just as in Figure 3, the individual stock book-to-market ratio R^2 series trends upwards as the training window expands. Across sample splits, not only is more time series data being used to train the procedure, but also the number of available individual stocks is increasing. Figure 4 shows (on the right-hand scale) that the number of stocks grows steadily from around 800 in 1955 to around 3600 in 1995. The R^2 for the PLS factor of firm-level book-to-market ratios increases more rapidly than other forecasters as the training sample is expands, suggesting that PLS forecasts learn relatively quickly as firms' book-to-market ratios become available and parameter estimates sharpen, rapidly overcoming the initial noise-induced R^2 deficit. By the mid-1980s the raw book-to-market R^2 series intersects that of *csp* and begins to provide better forecasts. However, it never reaches the high degree of predictability demonstrated

[23]Extracting PCs from the cross section of individual stock value ratios led to poorer results and so is not shown.

Figure 4: Out-of-Sample R^2 by Sample Split Date, One Month Returns



Notes: Out-of-sample R^2 by varying sample split date, given on the left scale. For forecasts of one month market returns from: a single PLS factor from the entire cross section of U.S. stocks' price-to-moving-average ratios (Individual Stock PMA) or book-to-market ratios (Individual Stock BM); a single PLS factor from 25 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993) (25 Fama-French Portfolio BMs); and alternative predictors taken from Goyal and Welch (2008) – the cross-section premium (csp) and the *ex post* consumption-wealth-income ratio (cay); and the first principal component extracted from the 25 portfolio book-to-market ratios (pc1). The number of stocks available in the cross-section is given on the right scale.

by our PLS approach applied to the cross section of firm-level valuation ratios constructed solely from stock prices. These results also join Polk et al. in highlighting the difficulty of working with noisy firm-level data.

III.B.5 Forecasting with Price-Dividend Ratios

As another robustness check, we consider the ability of a cross section of alternative valuation ratios to forecast market returns. In Section I, we note that the Campbell and Shiller (1988) present value identity produces a factor model for the cross section of log price-dividend ratios in direct analogy with Equation 6 under similar assumptions. Thus far, our analysis has focused on book-to-market ratios to avoid the lack of dividend payments

(and hence undefined price-dividend ratios) for a substantial fraction of U.S. firms.[24] While concerns about declining numbers of dividend-paying firms are partly mitigated by portfolio aggregation, in many cases portfolios can be dominated by non-dividend payers, resulting in an erratic and highly inflated price-dividend ratio for that portfolio.[25] In order to develop well-behaved portfolio price-dividend ratios, we form our own sets of six, 25 and 100 portfolios on the basis of underlying firms' market equity and book-to-market ratio, with the key difference that we exclude non-dividend paying firms. When forming portfolios, we only assign a stock to a portfolio in month t if at paid positive dividends in the twelve months prior to t.[26] This greatly increases the fraction of firms in our portfolios with well-defined price-dividend ratios, while continuing to condition portfolio formation only on past publicly available information. We refer to this sample as "past dividend payers." Dividend paying behavior is highly persistent among U.S. firms, so that a firm having paid dividends in the past twelve months strongly predicts that it will pay dividends in the subsequent twelve months.

Market return forecasts based on a single PLS factor extracted from the cross section of portfolio price-dividend ratios demonstrates strong predictive accuracy, on par with our results using book-to-market ratios. In-sample annual return R^2 numbers for six, 25 and 100 portfolios are 8.3%, 10.6% and 33.9%, and Kelly-Pruitt, Hodrick and Newey-West t-statistics are all significant at least at the 0.1% level. Out-of-sample R^2's are 13.7%, 4.8% and 9.6% respectively, all significant at least at the 1% level according to the Clark-McCracken test. The out-of-sample R^2 from 25 price-dividend ratios across various split dates behaves very similarly to those shown for 25 book-to-market ratios in Figure 3.

[24] The fraction of firms that paid dividends in 1946, 1980 and 2008 was 86%, 64% and 36%, though these fractions are substantially higher, 97%, 93% and 76%, when weighted by market capitalization.

[25] An earlier draft of this paper documents this behavior in detail. These results are available upon request.

[26] We use simultaneous two-way sorts, rebalance portfolios monthly, and most importantly, strictly preclude look-ahead information in portfolio construction, as is the case in the original Fama and French (1993) portfolios.

III.B.6 Forecasting Outside the U.S.

Our last robustness check asks whether our return predictability results hold internationally. To do so, we forecast returns on the value-weighted aggregate world portfolio (excluding the U.S.) by applying PLS to an international cross section of non-U.S., country-level valuation ratios. Monthly data for country-level portfolios are available from Ken French's website beginning in 1975, following the construction described in Fama and French (1998). This sample, based on data from MSCI, sorts equities from each country into a high value and low value portfolio. Countries covered in the sample are Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Italy, Japan, Malaysia, Netherlands, Norway, New Zealand, Singapore, Spain, Sweden, Switzerland and UK. We use cum- and ex-dividend returns on these portfolios to calculate price-dividend ratios of the two portfolios in each country, resulting in a cross section of 40 portfolios price-dividend ratios.[27] This cross section is used to forecast the return on an international equity index (excluding the U.S.), which is a value-weighted portfolio of country-level index returns in these 21 countries (portfolio weights are determined by a country's weight in the MSCI EAFE index). Because the data begin in 1975, we consider only monthly returns and take 1995 to be our out-of-sample split date.

We find that the world equity index return is highly predictable by country-level value ratios. The monthly out-of-sample R^2 is 2.3%, for which Clark and McCracken's (2001) test statistic is significant at least at the 1% level. Figure 5 shows that this strong out-of-sample performance is robust to a wide range of sample splits, and gradually increases with the length of the training sample as in the U.S. data. In-sample analysis produces an R^2 of 5.3%

[27] French's data includes price and dividend data at the monthly frequency, and aggregate country book-to-market ratios at the annual frequency. Because the sample begins, at the earliest, in 1975 (some countries have an even shorter sample), for meaningful out-of-sample analysis we conduct our forecasting analysis at the monthly frequency and therefore rely on price-dividend ratios rather than book-to-market ratios as our predictors. A comparison of annual price-dividend and book-to-market ratios suggests that the two series are highly similar at the country level. The median correlation between the two ratios is 86% (mean of 78%) across the 21 countries we study. In-sample forecasts of non-overlapping annual returns based on country-level book-to-market ratios generates an R^2 of 8.2%, versus an R^2 of 9.8% using price-dividend ratio data at the same frequency.

Figure 5: Out-of-Sample R^2 by Sample Split Data, One Month International Returns

Notes: Out-of-sample R^2 by varying sample split date. For forecasts of one month international stock returns from a single PLS factor from 42 book-to-market ratios of high- and low-value portfolios for 21 countries stocks from Fama and French (1998). See Section III.B.6 for the list of countries.

with a Kelly-Pruitt t-stat significant at least at the 0.1% level. The success of a single PLS factor drawn from the cross section of value ratios in an international sample lends further confidence to the robustness of our findings.

III.C Cash Flow Growth Predictions

Thus far we have focused on forecasts of aggregate market returns. Asset prices depend not only on discount rates, but also on expectations about assets' future cash flows. Hence it is important for our understanding of asset pricing to also investigate how much information valuation ratios contain about the market's expectations of future cash flow growth. The Vuolteenaho identity incorporates cash flow growth in terms of return on equity, while the Campbell-Shiller identity depends on dividend growth. Our analysis focuses on forecasting dividend growth and earnings growth, since these quantities have been at the center of growth forecasting in the asset pricing literature (see Ball and Watts (1972), Campbell and Shiller (1988), Cochrane (1992), Fama and French (2000), Lettau and Ludvigson (2005), Koijen and Van Nieuwerburgh (2010), and Lacerda and Santa-Clara (2010)).

Aggregate dividend growth is calculated from the universe of CRSP stocks and aggregate earnings growth data is calculated from Standard and Poor's data on Robert Shiller's website. Our analysis focuses on annual cash flow growth data in order to avoid spurious predictability arising from well-known within-year cash flow seasonality. Table V reports results from our PLS approach to forecasting annual aggregate U.S. dividend or earnings growth based on six, 25 and 100 Fama-French portfolio book-to-market ratios. Across all portfolio sorts and cash flow measures, the in-sample and out-of-sample results are positive and statistically significant. For dividend growth, the in-sample R^2 is between 21% and 44%, with an out-of-sample R^2 between 6% and 25%. Earnings growth forecast produce an in-sample R^2 between 9% and 29% with an out-of-sample R^2 around 3%.[28] Figure 6 plots the out-of-sample R^2 for dividend growth forecasts, comparing our procedure to forecasts based on the aggregate book-to-market ratio and the first principal component extracted from the cross section of book-to-market ratios. As we showed with returns, the strong out-of-sample predictive results from our procedure are robust across sample split and dominate aggregate value ratios and principal components. Figure 7 plots the in-sample and out-of-sample fitted dividend growth series alongside fits from aggregate valuation ratios. Not only do in-sample PLS estimates suggest a much more volatile series for conditional expected annual dividend growth, this is true out-of-sample as well, whose fits are nearly identical to in-sample estimates.

Only recently have more sophisticated econometric approaches begun to identify predictability in aggregate dividend growth, as in van Binsbergen and Koijen (2010). The cross section of book-to-market ratios identifies similarly large dividend growth forecastability, and we document the robustness of this fact out-of-sample. We also document new evidence of robust predictability in aggregate U.S. earning growth.

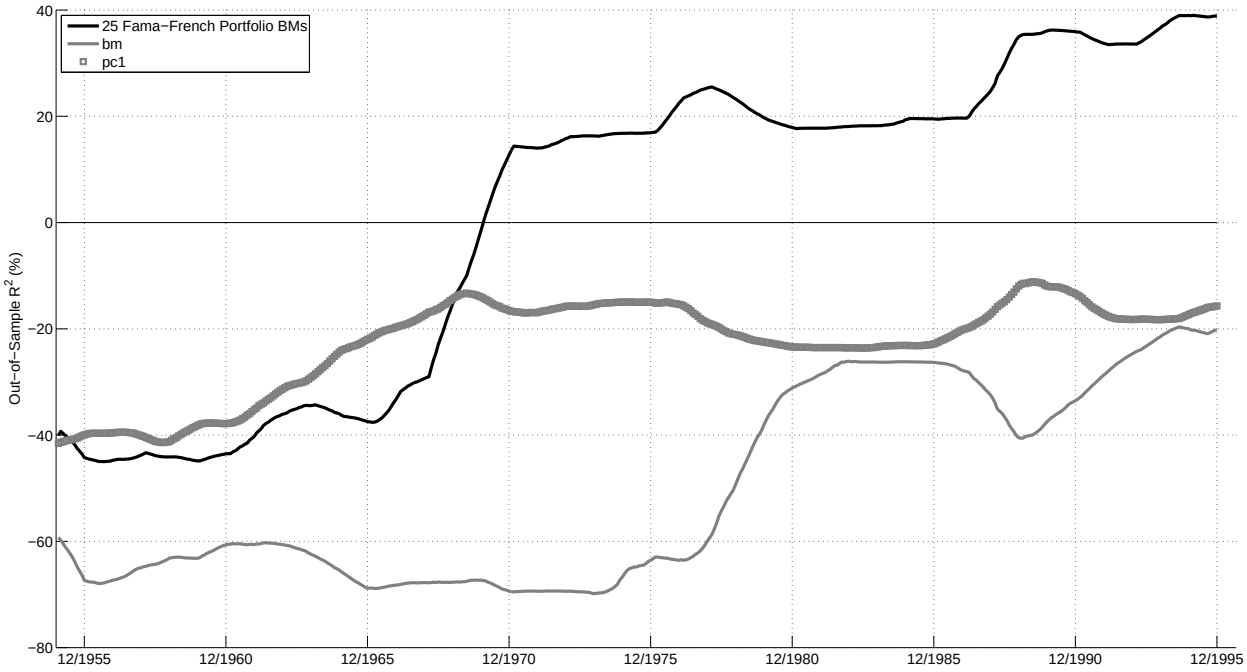
[28] Our annual dividend growth series is calculated assuming interim dividend payments are reinvested at the risk free rate. Alternative strategies that reinvest in the market portfolio convolute market return and dividend growth dynamics. We refer readers to Koijen and Van Nieuwerburgh (2010) for a detailed discussion of this point.

Table V: Market Cash Flow Predictions (1930-2009)

	R^2 (%)	p (KP/CM)	p (Hodrick)	p (NW)
Panel A: Dividend Growth				
6 Portfolios				
In-Sample	21.12	0.002	0.002	< 0.001
Out-of-Sample	25.44	< 0.010	-	-
25 Portfolios				
In-Sample	20.58	0.005	0.002	0.001
Out-of-Sample	19.53	< 0.010	-	-
100 Portfolios				
In-Sample	44.29	< 0.001	< 0.001	< 0.001
Out-of-Sample	6.13	< 0.010	-	-
Panel B: Earnings Growth				
6 Portfolios				
In-Sample	8.68	0.002	< 0.001	< 0.001
Out-of-Sample	2.76	< 0.100	-	-
25 Portfolios				
In-Sample	11.98	< 0.001	< 0.001	< 0.001
Out-of-Sample	2.63	< 0.050	-	-
100 Portfolios				
In-Sample	28.88	< 0.001	< 0.001	< 0.001
Out-of-Sample	3.49	< 0.050	-	-

Notes: Results of PLS forecasts of one year aggregate dividend or earnings growth for the U.S. stock market. The sets of predictor variables are six, 25 and 100 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993). In-sample results are for the 1930-2009 sample. Out-of-sample forecasts split the sample in 1980, using the pre-1980 period as a training window, and recursively forecasting returns beginning in January 1980 (results from a wide range of alternative sample splits are shown in Figure 3). We report in-sample and out-of-sample forecasting regression R^2 in percent. We also report p-values of three different in-sample test statistics. The first is the asymptotic predictive loading t-statistic from Kelly and Pruitt (2011), calculated on every non-overlapping set of residuals as described in the text, and also Hodrick (1992) and Newey-West (1985) t-statistic p-values. For out-of-sample tests we report Clark and McCracken's (2001) ENC-NEW encompassing test statistic in *italics* with p-values from Clark and McCracken's (2001) appendix tables, testing the null hypothesis of no forecast improvement over the historical mean and following Clark and McCracken (2005) in using Newey-West standard errors with twelve lags to consistently estimate the appropriate asymptotic variance.

Figure 6: Out-of-Sample R^2 by Sample Split Data, Annual Dividend Growth

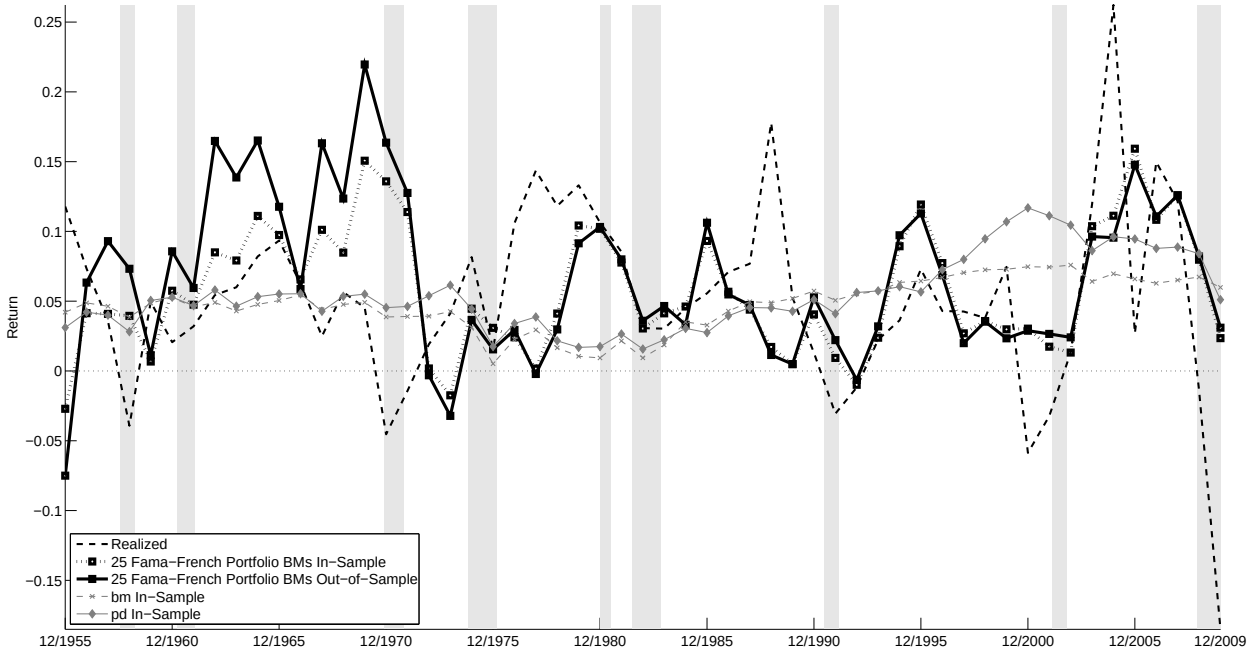


Notes: Out-of-sample R^2 by varying sample split date. For forecasts of annual aggregate dividend growth from a single PLS factor from 25 book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993) (25 Fama-French Portfolio BMs). We also plot results from forecasts based on the aggregate book-to-market ratio (bm) and the first principal component extracted from the 25 portfolio book-to-market ratios (pc1).

IV Conclusion

We derive a dynamic latent factor model representation for the cross section of asset valuation ratios. The same factors that drive these present values also determine aggregate expectations of market returns and cash flow growth, enabling us to make use rich cross-sectional information in constructing forecasts. To analyze these latent processes, we use the method of partial least square and recent econometric results on its behavior in factor model settings. By extracting information from disaggregate valuation ratios we are able to construct remarkably accurate forecasts of returns and cash flow growth rates both in-sample and out-of-sample. The resulting estimates reveal important facets of the time series of expected returns and cash flow growth that may be used to guide future models of cash flows and discount rates. In particular, market expectations are much more volatile and less

Figure 7: Dividend Growth Predictions



Notes: The figure shows year-end (December) realized dividend growth for the aggregate U.S. stock market (Realized), both in-sample (PLS IS) and out-of-sample (PLS OOS) forecasts of aggregate growth from a univariate PLS factor extracted from 25 Fama-French book-to-market ratios of size and value-sorted portfolios of U.S. stocks from Fama and French (1993), and in-sample forecasts from predictive regressions on the aggregate book-to-market (bm IS) and aggregate price-dividend (pd IS) ratios. NBER recession dates are represented by the shaded area.

autocorrelated than previous results showed. Our results are robust across a variety of cross sections, out-of-sample procedures and hold in both U.S. and international data. The cross section of valuation ratios, as present value identities imply, hold a wealth of information about investor expectations. A deeper understanding of economic fundamentals that drive these valuation ratios is a promising avenue for future research.

References

Ang, A., and G. Bekaert. 2007. "Stock Return Predictability: Is it There?" *Review of Financial Studies*, 20(3): 651.

Bai, Jushan, and Serena Ng. 2006. "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions." *Econometrica*, 74(4): 1133–1150.

Ball, R., and R. Watts. 1972. "Some time series properties of accounting income." *The Journal of Finance*, 27(3): 663–681.

Bansal, R., and S. Viswanathan. 1993. "No Arbitrage and Arbitrage Pricing: A New Approach." *Journal of Finance*, 48(4): 1231–1262.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 59(4): 1481–1509.

Bansal, Ravi, Robert Dittmar, and Christian Lundblad. 2005. "Consumption, Dividends, and the Cross Section of Equity Returns." *Journal of Finance*, 60(4): 1639–1672.

Binsbergen, Jules, and Ralph Koijen. 2010. "Predictive Regressions: A Present-Value Approach." *Journal of Finance*, forthcoming.

Brennan, M.J., and Y. Xia. 2005. "Persistence, Predictability, and Portfolio Planning." *Working Paper, Rodney L. White Center*.

Campbell, John, and John Cochrane. 1999. "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107(2): 205–251.

Campbell, John, and Robert Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1(3): 195–228.

Campbell, John, and Samuel Thompson. 2008. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, 21(4): 1509–1531.

Chowdhry, Bhagwan, Richard Roll, and Yihong Xia. 2005. "Extracting Inflation from Stock Returns to Test Purchasing Power Parity." *American Economic Review*, 95(1): 255–276.

Clark, Todd, and Michael McCracken. 2001. "Tests of Equal Forecast Accuracy and Encompassing for Nested Models." *Journal of Econometrics*, 105(1): 85–110.

Clark, Todd, and Michael McCracken. 2005. "Evaluating Direct Multistep Forecasts." *Econometric Reviews*, 24(4): 369–404.

Cochrane, J.H. 1992. "Explaining the variance of price-dividend ratios." *Review of Financial Studies*, 5(2): 243–80.

Cochrane, J.H. 1999. "New Facts in Finance." *Economic Perspectives*, 23.

Cochrane, John. 2005. *Asset Pricing.* Princeton University Press.

Cochrane, John. 2008*a*. "State-Space vs. VAR Models for Stock Returns." *Working Paper, Chicago Booth.*

Cochrane, John. 2008*b*. "The Dog That Did Not Bark: A Defense of Return Predictability." *Review of Financial Studies*, 21(4): 1533.

Connor, G. 1984. "A Unified Beta Pricing Theory." *Journal of Economic Theory*, 34(1): 13–31.

Fama, E.F., and K.R. French. 2000. "Forecasting Profitability and Earnings." *The Journal of Business*, 73(2): 161–175.

Fama, Eugene, and James MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *The Journal of Political Economy*, 81(3): 607–636.

Fama, Eugene, and Kenneth French. 1988. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy*, 96(2): 246–73.

Fama, Eugene, and Kenneth French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33(1): 3 – 56.

Fama, Eugene F., and Kenneth R. French. 1998. "Value versus Growth: The International Evidence." *Journal of Finance*, 53(6): 1975–1999.

Ferreira, M., and P. Santa Clara. 2010. "Forecasting Stock Market Returns: The Sum of the Parts Is More Than the Whole." *Journal of Financial Economics*, forthcoming.

Ghosh, A., and G. Constantinides. 2010. "The Predictability of Returns with Regime Shifts in Consumption and Dividend Growth." *Working Paper, Chicago Booth.*

Goetzmann, W.N., and P. Jorion. 1993. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, 48(2): 663–679.

Goyal, A., and I. Welch. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies*, 21(4): 1455–1508.

Hansen, L.P., J.C. Heaton, and N. Li. 2008. "Consumption Strikes Back? Measuring Long-Run Risk." *Journal of Political Economy*, 116(2).

Hansen, Peter Reinhard, and Allan Timmermann. 2011. "Choice of Sample Split in Out-of-Sample Forecast Evaluation." Stanford University.

Hodrick, R.J. 1992. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial studies*, 5(3): 357.

Huber, Peter J. 1973. "Robust Regression: Asymptotics, Conjectures and Monte Carlo." *Annals of Statistics*, 1(5): 799–821.

Inoue, Atsushi, and Barbara Rossi. 2011. "Out-of-Sample Forecast Tests Robust to the Choice of Window Size." C.E.P.R. Discussion Papers 8542.

Jagannathan, Ravi, and Zhenyu Wang. 1998. "A Note on the Asymptotic Covariance in Fama-MacBeth Regression." *Journal of Finance*, 53(2): 799–801.

Kelly, B. 2011. "Tail Risk and Asset Prices." *Working Paper, Chicago Booth*.

Kelly, B., and S. Pruitt. 2011. "The Three-Pass Regression Filter: A New Approach to Forecasting with Many Predictors." *Working Paper, Chicago Booth*.

Kiku, D. 2006. "Is the Value Premium a Puzzle?" *Working Paper, The Wharton School*.

Koijen, R., and S. Van Nieuwerburgh. 2010. "Predictability of Returns and Cash Flows." *Working Paper, Chicago Booth*.

Kothari, SP, and J. Shanken. 1997. "Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis." *Journal of Financial Economics*, 44(2): 169–203.

Lacerda, F., and P. Santa-Clara. 2010. "Foreasting dividend growth to better predict returns." *Working Paper*.

Lettau, M., and S. Van Nieuwerburgh. 2008. "Reconciling the Return Predictability Evidence." *Review of Financial Studies*, 1607–1652.

Lettau, M., and S.C. Ludvigson. 2005*a*. "Expected returns and expected dividend growth." *Journal of Financial Economics*, 76(3): 583–626.

Lettau, Martin, and Sydney Ludvigson. 2005*b*. "Expected Returns and Expected Dividend Growth." *Journal of Financial Economics*, 76(3): 583–626.

Lewellen, J. 2004. "Predicting Returns with Financial Ratios." *Journal of Financial Economics*, 74(2): 209–235.

Lintner, J. 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *The Review of Economics and Statistics*, 47(1): 13–37.

Ludvigson, Sydney C., and Serena Ng. 2007. "The empirical risk-return relation: A factor analysis approach." *Journal of Financial Economics*, 83(1): 171–222.

Menzly, L., T. Santos, and P. Veronesi. 2004. "Understanding Predictability." *Journal of Political Economy*, 112(1): 1–47.

Merton, R.C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica*, 41(5): 867–887.

Miller, Merton H, and Myron S Scholes. 1982. "Dividends and Taxes: Some Empirical Evidence." *Journal of Political Economy*, 90(6): 1118–41.

Nelson, C.R., and M.J. Kim. 1993. "Predictable Stock Returns: The Role of Small Sample Bias." *Journal of Finance*, 48(2): 641–661.

Newey, Whitney K, and Kenneth D West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3): 703–08.

Pástor, L., and P. Veronesi. 2006. "Was There a Nasdaq Bubble in the Late 1990s?" *Journal of Financial Economics*, 81(1): 61–100.

Pástor, Lubos, and Pietro Veronesi. 2003. "Stock Valuation and Learning about Profitability." *Journal of Finance*, 58(5): 1749–1790.

Pástor, Lubos, Meenakshi Sinha, and Bhaskaran Swaminathan. 2008. "Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital." *Journal of Finance*, 63(6): 2859–2897.

Paye, B.S., and A. Timmermann. 2006. "Instability of Return Prediction Models." *Journal of Empirical Finance*, 13(3): 274–315.

Polk, C., S. Thompson, and T. Vuolteenaho. 2006. "Cross-Sectional Forecasts of the Equity Premium." *Journal of Financial Economics*, 81(1): 101–141.

Rozeff, M.S. 1984. "Dividend Yields Are Equity Risk Premiums." *Journal of Portfolio Management*, 11(1): 68–75.

Rytchkov, O. 2008. "Filtering Out Expected Dividends and Expected Returns." *Working Paper, Temple University*.

Shanken, Jay. 1992. "The Current State of the Arbitrage Pricing Theory." *Journal of Finance*, 47(4): 1569–74.

Sharpe, W.F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk." *Journal of Finance*, 19(3): 425–442.

Stambaugh, R. 1986. "Bias in Regressions with Lagged Stochastic Regressors." *Working Paper, University of Chicago*.

Stambaugh, R.F. 1999. "Predictive Regressions." *Journal of Financial Economics*, 54(3): 375–421.

Stock, James H., and Mark W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association*, 97(460): 1167–1179.

Stock, James H., and Mark W. Watson. 2011. "Generalized Shrinkage Methods for Forecasting Using Many Predictors." Princeton University Working Papers.

Treynor, J. 1961. "Toward a Theory of Market Value of Risky Assets." *Unpublished manuscript*.

Vuolteenaho, Tuomo. 2002. "What Drives Firm-Level Stock Returns?" *Journal of Finance*, 57(1): 233–264.

Wold, Harald. 1975. "Estimation of principal components and related models by iterative least squares." In *Multivariate Analysis.* , ed. P.R. Krishnaiaah, 391–420. New York: Academic Press.

A Appendix

A.A Derivation of Present Value System

$$
\begin{aligned}
v_{i,t} &= \frac{\kappa_i}{1-\rho_i} + \sum_{j=0}^{\infty} \rho_i^j (-r_{i,t+j+1} + \Delta cf_{i,t+j+1}) \\
&= \frac{\kappa_i}{1-\rho_i} + \sum_{j=0}^{\infty} \rho_i^j \mathbb{E}_t(-\mu_{i,t+j} + g_{i,t+j}) \\
&= \frac{\kappa_i}{1-\rho_i} + \sum_{j=0}^{\infty} \rho_i^j \mathbb{E}_t \left[-(\gamma_{i,0} + \boldsymbol{\gamma}_i' \boldsymbol{F}_{t+j}) + (\delta_{i,0} + \boldsymbol{\delta}_i' \boldsymbol{F}_{t+j}) + \varepsilon_{i,t+j} \right] \\
&= \frac{\kappa_i - \gamma_{i,0} + \delta_{i,0}}{1-\rho_i} + \sum_{j=0}^{\infty} \rho_i^j \left(\mathbb{E}_t [\boldsymbol{\iota}' \boldsymbol{\Upsilon}_i' \boldsymbol{F}_{t+j} + \varepsilon_{i,t+j}] \right) \\
&= \frac{\kappa_i - \gamma_{i,0} + \delta_{i,0}}{1-\rho_i} + \sum_{j=0}^{\infty} \rho_i^j \boldsymbol{\iota} \boldsymbol{\Upsilon}_i' \boldsymbol{\Lambda}_1^j \boldsymbol{F}_t + \varepsilon_{i,t} \\
&= \frac{1}{1-\rho_i} (\kappa_i - \gamma_{i,0} + \delta_{i,0}) + \boldsymbol{\iota}' \boldsymbol{\Upsilon}_i' (I - \rho_i \boldsymbol{\Lambda}_1)^{-1} \boldsymbol{F}_t + \varepsilon_{i,t} \\
&= \phi_{i,0} + \boldsymbol{\phi}_i' \boldsymbol{F}_t + \varepsilon_{i,t}
\end{aligned}
$$

where we have defined $\boldsymbol{\Upsilon}_i = (\boldsymbol{\gamma}_i, \boldsymbol{\delta}_i)$, $\boldsymbol{\iota} = (-1, 1)'$, $\boldsymbol{\phi}_i' = \boldsymbol{\iota}' \boldsymbol{\Upsilon}_i' (I - \rho_i \boldsymbol{\Lambda}_1)^{-1}$, and $\phi_{i,0} = \frac{1}{1-\rho_i} (\kappa_i - \gamma_{i,0} + \delta_{i,0})$.

A.B Partial Least Squares Assumptions

We adapt the assumptions of Kelly and Pruitt (2011) for their analysis of the three-pass regression filter, a generalization of partial least squares. That paper has extensive discussion of the assumptions: Suffice it to say that they are adequately weak and are satisfied by our model here. There is a final condition in that paper that is trivially satisfied by our application because we use the forecast target itself as our proxy variable in the first-pass regressions.

Assumption 1 (Factor Structure)**.** *The data are generated by the following:*

$$
\begin{aligned}
\boldsymbol{v}_t &= \boldsymbol{\phi}_0 + \boldsymbol{\Phi} \boldsymbol{F}_t + \boldsymbol{\varepsilon}_t & \qquad y_{t+h} &= \beta_0 + \boldsymbol{\beta}' \boldsymbol{F}_t + \eta_{t+h} \\
\boldsymbol{V} &= \boldsymbol{\iota} \boldsymbol{\phi}_0' + \boldsymbol{F} \boldsymbol{\Phi}' + \boldsymbol{\varepsilon} & \qquad \boldsymbol{y} &= \boldsymbol{\iota} \beta_0 + \boldsymbol{F} \boldsymbol{\beta} + \boldsymbol{\eta}
\end{aligned}
$$

where $\boldsymbol{F}_t = (\boldsymbol{f}_t', \boldsymbol{g}_t')'$, $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_f, \boldsymbol{\Phi}_g)$, *and* $\boldsymbol{\beta} = (\boldsymbol{\beta}_f', \boldsymbol{0}')'$ *with* $|\boldsymbol{\beta}_f| > \boldsymbol{0}$. $K_f > 0$ *is the dimension of vector* $\boldsymbol{f}_t$, $K_g \geq 0$ *is the dimension of vector* $\boldsymbol{g}_t$, *and* $K = K_f + K_g$.

Assumption 2 (Factors, Loadings and Residuals)**.** *Let* $M < \infty$. *For any* i, s, t

1. $\mathbb{E}\|\boldsymbol{F}_t\|^4 < M$, $T^{-1} \sum_{s=1}^{T} \boldsymbol{F}_s \xrightarrow[T\to\infty]{p} \boldsymbol{\mu}$ *and* $T^{-1} \boldsymbol{F}' \boldsymbol{J}_T \boldsymbol{F} \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_F$
2. $\mathbb{E}\|\boldsymbol{\phi}_i\|^4 \leq M$, $N^{-1} \sum_{j=1}^{N} \boldsymbol{\phi}_j \xrightarrow[T\to\infty]{p} \bar{\boldsymbol{\phi}}$, $N^{-1} \boldsymbol{\Phi}' \boldsymbol{J}_N \boldsymbol{\Phi} \xrightarrow[N\to\infty]{p} \boldsymbol{\mathcal{P}}$ *and* $N^{-1} \boldsymbol{\Phi}' \boldsymbol{J}_N \boldsymbol{\phi}_0 \xrightarrow[N\to\infty]{p} \boldsymbol{P}_1$[29]
3. $\mathbb{E}(\varepsilon_{it}) = 0, \mathbb{E}|\varepsilon_{it}|^8 \leq M$
4. $\mathbb{E}(\boldsymbol{\omega}_t) = \boldsymbol{0}, \mathbb{E}||\boldsymbol{\omega}_t||^4 \leq M, T^{-1/2} \sum_{s=1}^{T} \boldsymbol{\omega}_s = \boldsymbol{O}_p(1)$ *and* $T^{-1} \boldsymbol{\omega}' \boldsymbol{J}_T \boldsymbol{\omega} \xrightarrow[N\to\infty]{p} \boldsymbol{\Delta}_\omega$
5. $\mathbb{E}_t(\eta_{t+h}) = \mathbb{E}(\eta_{t+h} | y_t, F_t, y_{t-1}, F_{t-1}, ...) = 0$, $\mathbb{E}(\eta_{t+h}^2) = \delta_\eta < \infty$ *for any* $h > 0$, *and* η_t *is independent of* $\phi_i(m)$ *and* $\varepsilon_{i,s}$.

[29] $\|\boldsymbol{\phi}_i\| \leq M$ can replace $\mathbb{E}\|\boldsymbol{\phi}_i\|^4 \leq M$ if $\boldsymbol{\phi}_i$ is non-stochastic.

Assumption 3 (Dependence)**.** *Let $x(m)$ denote the m^{th} element of $\boldsymbol{x}$. For $M < \infty$ and any i, j, t, s, m_1, m_2*

1. $\mathbb{E}(\varepsilon_{it}\varepsilon_{js}) = \sigma_{ij,ts}$, $|\sigma_{ij,ts}| \leq \bar{\sigma}_{ij}$ *and* $|\sigma_{ij,ts}| \leq \tau_{ts}$, *and*
   - *(a)* $N^{-1}\sum_{i,j=1}^{N} \bar{\sigma}_{ij} \leq M$
   - *(b)* $T^{-1}\sum_{t,s=1}^{T} \tau_{ts} \leq M$
   - *(c)* $N^{-1}\sum_{i,s} |\sigma_{ii,ts}| \leq M$
   - *(d)* $N^{-1}T^{-1}\sum_{i,j,t,s} |\sigma_{ij,ts}| \leq M$
2. $\mathbb{E}\left|N^{-1/2}T^{-1/2}\sum_{s=1}^{T}\sum_{i=1}^{N}\left[\varepsilon_{is}\varepsilon_{it} - \mathbb{E}\left(\varepsilon_{is}\varepsilon_{it}\right)\right]\right|^2 \leq M$
3. $\mathbb{E}\left|T^{-1/2}\sum_{t=1}^{T} F_t(m_1)\omega_t(m_2)\right|^2 \leq M$
4. $\mathbb{E}\left|T^{-1/2}\sum_{t=1}^{T} \omega_t(m_1)\varepsilon_{it}\right|^2 \leq M$.

Assumption 4 (Central Limit Theorems)**.** *For any i, t*

1. $N^{-1/2}\sum_{i=1}^{N} \boldsymbol{\phi}_i\varepsilon_{it} \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{\Gamma}_{\Phi\varepsilon}\right)$, *where* $\boldsymbol{\Gamma}_{\Phi\varepsilon} = \text{plim}_{N\to\infty} N^{-1}\sum_{i,j=1}^{N}\mathbb{E}\left[\boldsymbol{\phi}_i\boldsymbol{\phi}_j'\varepsilon_{it}\varepsilon_{jt}\right]$
2. $T^{-1/2}\sum_{t=1}^{T} \boldsymbol{F}_t\eta_{t+h} \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{\Gamma}_{F\eta}\right)$, *where* $\boldsymbol{\Gamma}_{F\eta} = \text{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T}\mathbb{E}\left[\eta_{t+h}^2\boldsymbol{F}_t\boldsymbol{F}_t'\right] > 0$
3. $T^{-1/2}\sum_{t=1}^{T} \boldsymbol{F}_t\varepsilon_{it} \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{\Gamma}_{F\varepsilon,i}\right)$, *where* $\boldsymbol{\Gamma}_{F\varepsilon,i} = \text{plim}_{T\to\infty} T^{-1}\sum_{t,s=1}^{T}\mathbb{E}\left[\boldsymbol{F}_t\boldsymbol{F}_s'\varepsilon_{it}\varepsilon_{is}\right] > 0$.

Assumption 5 (Normalization)**.** $\boldsymbol{\mathcal{P}} = \boldsymbol{I}$, $\boldsymbol{P}_1 = \boldsymbol{0}$ *and* $\boldsymbol{\Delta}_F$ *is diagonal, positive definite, and each diagonal element is unique.*