

Labor Income and Predictable Stock Returns

Tano Santos

Columbia University and NBER

Pietro Veronesi

University of Chicago, NBER, and CEPR

We propose a novel economic mechanism that generates stock return predictability in both the time series and the cross-section. Investors' income has two sources, wages and dividends that grow stochastically over time. As a consequence the fraction of total income produced by wages fluctuates depending on economic conditions. We show that the risk premium that investors require to hold stocks varies with these fluctuations. A regression of stock returns on lagged values of the labor income to consumption ratio produces statistically significant coefficients and large adjusted R^2 s. Tests of the model's cross-sectional predictions on the set of 25 Fama–French portfolios sorted on size and book-to-market are also met with considerable support.

Researchers, at least since Mayers (1972), have long recognized the importance of accounting for labor income, and, more generally, human capital, in asset pricing tests. Indeed, labor income constitutes around 75% of consumption and human capital is a significant component of wealth. In this article, we propose a minimal extension of the standard consumption asset pricing model, where consumption is funded by sources other than financial, labor income in particular, and that allows for tractable and interpretable formulas for prices and returns. The model shows that allowing for this alternative source of consumption immediately yields implications for the dynamics of asset prices. In particular, we show, theoretically and empirically, that fluctuations in the fraction of consumption funded by labor income results in stock return predictability both in the time series and the cross-section.

To illustrate the role of labor income on the predictability of stock returns, we consider a general equilibrium model, where a representative agent receives income from two sources, financial and nonfinancial (human), and where the mix between the two sources of income varies over time. We first show that changes in the fraction of consumption

We thank Fernando Alvarez, Nick Barberis, John Y. Campbell, John H. Cochrane, George Constantinides, Kent Daniel, Gene Fama, Lars P. Hansen, John Heaton, Michael Johannes, Martin Lettau, Sydney Ludvigson, Lubos Pastor, Zhenyu Wang, and seminar participants at the 2000 NBER Summer Institute in Asset Pricing, Columbia University, Northwestern University, and the University of Chicago. We thank Lior Menzly for excellent research assistance. All errors are our own. Address correspondence to Pietro Veronesi, University of Chicago GSB, 5807 South Woodlawn Avenue, Chicago IL 60637, or e-mail: fpverone@gsb.uchicago.edu.

funded by labor income induce fluctuations in the expected excess return of the market portfolio. The intuition for this result is straightforward. If, say, most of consumption is funded by labor income, financial assets constitute a small fraction of consumption and thus co-vary little with it. For this reason, investors require a low premium to hold them. It follows then that the ratio of labor income to consumption should forecast stock returns at the aggregate level. Notice that the time-series predictability of aggregate returns stems solely from fluctuations in the fraction of consumption funded by labor income. No other ingredient is required to generate this predictability, whether it be habit persistence, as in Campbell and Cochrane (1999), house money effects as in Barberis, Huang, and Santos (2001), or learning effects as in Timmermann (1993) and Veronesi (2000). Our model then provides an alternative source of predictability so far unexplored in the literature.

We test this implication by regressing aggregate market returns on the labor income to consumption ratio and find that this variable is a strong predictor of long-horizon returns. This result is robust to alternative constructions of this ratio and the inclusion of the dividend yield. In addition, when running a similar regression in model-simulated data, the share of labor income to consumption is also a significant predictor of long-horizon returns.

The ability of the labor income to consumption ratio to forecast long-horizon returns suggests that this variable is also a useful predictor of the cross-section. The failure of the capital asset pricing model (CAPM) to explain the cross-section of stock returns stands as a central finding in empirical asset pricing, and many have argued for a role for labor income in tests of the cross-section. For instance, a common concern, famously emphasized by Roll (1977), is that the return on the value-weighted portfolio of assets listed in major U.S. stock exchanges is typically taken as a proxy for the return on the market portfolio and that this proxy may not be good enough. Building on this insight, Campbell (1996) and Jagannathan and Wang (1996) include a measure of the return on human capital as part of the returns on aggregate wealth. Recently, as well, Lettau and Ludvigson (2001a, 2001b) have shown that a variable that measures deviations of consumption from its stable relation with wealth, that includes both human and nonhuman (financial) wealth, has remarkable predictive power both in the time series and the cross-section of stock returns.

In our framework, the CAPM with respect to the total wealth portfolio holds conditionally if agents have log utility. In general, although the conditional CAPM does not hold for relative risk aversion different than one, we show that it holds approximately. We obtain the asset's beta in closed form for the log utility case and show that it is a function of two variables. The first one is related to the current size of the asset's

contribution to consumption. If this contribution is significant and covaries positively with consumption growth, the asset will command a higher premium than an otherwise identical asset with a lower contribution to consumption. The second variable is, again, the labor income to consumption ratio that captures changes in the overall covariance between the return on the total wealth portfolio and financial assets. It follows then that taking into account labor income in asset pricing requires more than updating the definition of the “market” portfolio (the total wealth portfolio.) It requires the use of conditioning information as well.

In addition, our model also allows us to assess the effect that sorting procedures based on prices have on tests of the cross-section. In our setup, sorting by the price of the security, normalized by dividends, is akin to sorting by expected dividend growth, which is the source of cross-sectional differences in expected excess returns. Still, we show that the sorting per se is not enough to capture the cross-sectional dispersion of average returns and that conditioning by the labor income to consumption ratio is key to fully capturing this dispersion.

We test these predictions in the set of 25 portfolios sorted by size and book-to-market introduced by Fama and French (1993) and find that the conditional CAPM performs considerably better than its unconditional counterpart when using the labor income to consumption ratio as a conditioning variable. We also run extensive simulations of the model to show that it can reproduce both the poor performance of the unconditional CAPM and the better one of its conditional version. To do so, we simulate financial data for multiple assets which we then sort into portfolios according to their price–dividend ratios. We then test the unconditional and the conditional CAPM in this set of fictitious test portfolios. Our simulations show that the model can reproduce the flat relation between the average and the excess returns predicted by the unconditional CAPM as well as the significant pricing errors that are usual in tests of the CAPM. Instead, in our conditional specification, fitted and average returns line up nicely along the 45° ray and the pricing errors are not significantly different than 0. In our framework then the value-spread puzzle is less of a puzzle.

Our model is related to a number of recent papers that attempt to model multiple securities in a general equilibrium setting.¹ The paper closest to ours in what concerns the cash-flow model is Menzly, Santos, and Veronesi (2004, MSV henceforth). However, there are several key differences with respect to that paper. First, the cash-flow model proposed here is slightly more general than the one in MSV. Our pricing formulas then apply to this more general class of cash-flow models.

¹ See Cochrane, Longstaff, and Santa-Clara (2004), Longstaff and Piazzesi (2004), and Menzly, Santos, and Veronesi (2004).

Second, the economic mechanism at work in each of these two models is very different; while MSV focuses on the interaction between time-varying risk preferences and fluctuations in expected dividend growth to explain the relation between valuation ratios and asset returns, this article has constant risk preferences. By construction, we shut down in this model the main source of variation in MSV, and thus the two models are exact opposites in terms of the economic mechanism that matter for changes in expected returns. As already mentioned, we focus here on the variation in the labor income/financial income mix to explain the variation over time of asset returns. Third, and finally, MSV focuses its empirical exercises on industry portfolios, whereas here we concentrate on a set of price-sorted portfolios, such as the size and book-to-market sorted portfolios. Our model offers a framework where the sorting procedure can be easily interpreted and where the role of labor income in pricing this particular set of test portfolios can be cleanly investigated.

Our article places itself at the intersection of two strands of the literature on asset pricing. On the one hand, the article contributes to the body of work that documents the time-series predictability of the aggregate market returns. The early predictability literature documents the forecasting power of prices scaled by either dividends or earnings and of various interest rate measures.² More recently, Lettau, and Ludvigson (2001a) manipulate the budget constraint to show that the consumption to total wealth ratio contains information about stock returns. Our article adds to this literature by providing yet more evidence on the predictability of stock returns. A critical difference between our work and previous empirical research though is that our predictive variable is neither a version of the stock price scaled by either dividends or earnings nor some other financial variable like the term premium, but rather a pure macroeconomic variable. Furthermore, given that it is directly observable, it does not need to be estimated, as in Lettau and Ludvigson (2001a, 2001b). Finally, it is important to emphasize that our testable implication does not result from basic manipulations of either the definition of returns [Cochrane (2001), p. 395–396] or the budget constraint.

The present article also adds to the growing body of empirical research that concentrates on testing the conditional versions of the CAPM. For instance, researchers like Cochrane (1996), Ferson and Harvey (1999), Jagannathan and Wang (1996), and Lettau and Ludvigson (2001b) use “conditioned down” versions of the CAPM (or the consumption CAPM) to obtain improvements in the ability of these models to describe the cross-section of stock returns. Others, like Bollerslev, Engle, and

² Campbell and Shiller (1988a, 1988b), Fama and French (1988a, 1988b), Hodrick (1992), and Lamont (1998) document the predictive power of prices scaled by dividends and earnings. Campbell (1987), Fama and French (1989), Hodrick (1992), and Keim and Stambaugh (1986) show the forecasting power of interest rate measures.

Wooldridge (1988), use GARCH methods to explore the role of changing betas in the context of the conditional CAPM. They find that the conditional covariances of each return with the market return vary substantially and are a significant determinant of time-varying risk premiums. Finally, Fama and French (1997) and Ferson and Harvey (1991), among others, parameterize both market risk premiums and factor loadings as functions of both aggregate variables and individual asset characteristics. They find substantial variation in both premiums and loadings.³ Few of these papers, though, derive the expression of beta from a full-fledged general equilibrium model and identify the variables that should proxy for the investors' information set from theoretical considerations. It is then one of the objectives of this article to place tests of conditional models on firmer theoretical ground.

In summary, our model hopes to provide a coherent view of the time-series and cross-sectional predictability of returns. Indeed, since Merton (1973), it has been understood that variables that predict market returns are natural conditioning variables for tests of the cross-section.⁴ Our contribution here is to show how taking into account labor income generates these two forms of predictability and to clarify the link between them. In addition, and as Lettau and Ludvigson (2004) have recently emphasized, these findings highlight the fact that information contained in consumption and labor income may be relevant for the long run valuation of financial assets.

The rest of the article is organized as follows. Section 1 presents the model. Section 2 derives the implications of the model for prices and returns. Section 3 presents the empirical results and Section 4 concludes.

1. The Model

Preferences. We assume the existence of a representative consumer whose preferences over aggregate consumption C_t are represented by the instantaneous utility function

³ The literature on the variation in betas is large. See Bollerslev, Engle, and Wooldridge (1988), Bodurtha and Mark (1991), Braun, Nelson, and Sunier (1995), Cho and Engle (1999), Evans (1994), and Ng (1991). Ball and Kothari (1989) argue that variation in betas and market premiums can explain the performance of winners and losers. Campbell (1987), Ferson (1989), Ferson, Kandel, and Stambaugh (1987), Harvey (1989), and Shanken (1990) develop tests of asset-pricing models with time-varying risk premiums and betas. Franzoni (2001) uses rolling regressions to document changes in the betas of growth and value stocks. For a criticism of the ability of time variation in betas to address puzzles in the cross-section, see Lewellen and Nagel (2003).

⁴ See Cochrane (1996), Ferson and Harvey (1999), and Lettau and Ludvigson (2001b) for recent contributions in this direction. For example, Ferson and Harvey (1999) state that "simple proxies for time variation in expected returns, based on common lagged instruments, are also significant cross-sectional predictors of returns."

$$U(C_t, t) = \begin{cases} \frac{e^{-\phi t} C_t^{1-\gamma}}{(1-\gamma)} & \text{if } \gamma \neq 1 \\ e^{-\phi t} \log(C_t) & \text{if } \gamma = 1 \end{cases}, \quad (1)$$

where γ is the coefficient of relative risk aversion and ϕ , the subjective discount rate.

Endowment. We assume that consumption is funded by labor income, w_t , and the proceeds from an investment in $n - 1$ additional financial securities, whose instantaneous dividend streams we denote by D_t^i , for $i = 2, \dots, n$. For notational convenience, let $D_t^1 = w_t$.

Modeling choices concerning $\{D_t^i\}_{i=1}^n$ and C_t cannot be made independently as market clearing imposes $C_t = w_t + \sum_{i=2}^n D_t^i$. Thus, particular assumptions on the processes governing $\{D_t^i\}_{i=1}^n$ induce in turn a specific aggregate consumption process. This is the essence of the difficulty when studying models with multiple assets, namely, how to model dividends and consumption so that they are empirically plausible, mutually consistent, and, at the same time, tractable enough to yield interpretable formulas for prices and returns.⁵ It is useful then to briefly review the nature of these difficulties to better motivate our specific assumptions.

Let $\mathbf{D}_t = (D_t^1, \dots, D_t^n)$ be the vector of dividends. Assume for instance that dividends D_t^i are given by

$$\frac{dD_t^i}{D_t^i} = \mu_D^i(\mathbf{D}_t)dt + \mathbf{v}_i' d\mathbf{B}_t \quad (2)$$

for some drifts $\mu_D^i(\mathbf{D}_t)$, where \mathbf{v}_i is an $n \times 1$ constant vector and $d\mathbf{B}_t$ is an $n \times 1$ vector of Brownian motions. The process for aggregate consumption C_t can now be written as

$$\frac{dC_t}{C_t} = \mu_c(\mathbf{s}_t)dt + \boldsymbol{\sigma}_c(\mathbf{s}_t)' d\mathbf{B}_t,$$

where $\mathbf{s}_t = (s_t^1, \dots, s_t^n)' = (D_t^1/C_t, \dots, D_t^n/C_t)'$ are the share of consumption produced by dividends (or labor income, for $i = 1$), and

$$\mu_c(\mathbf{s}_t) = \sum_{i=1}^n s_t^i \mu_D^i \quad (3)$$

$$\boldsymbol{\sigma}_c(\mathbf{s}_t) = \sum_{i=1}^n s_t^i \mathbf{v}_i. \quad (4)$$

⁵ See, also, Bossaerts and Green (1989). Cochrane, Longstaff and Santa-Clara (2004) were recently able to obtain closed form formulas for prices in the special case where there are only two assets whose dividends are log-normal, and agents have log utility. Their model, however, implies that in the long run, one of the two assets would dominate the economy with probability one.

One difficulty in obtaining tractable and interpretable formulas for prices is the dependence of the drift and volatility of the consumption process on the shares $\mathbf{s}_t = (s_t^1, \dots, s_t^n)'$. Still, by making judicious but economically motivated choices for both the consumption and share processes, closed form solutions can be obtained for both prices and returns. The next two assumptions summarize the essence of our cash-flow model.

Assumption 1. *The aggregate consumption process is given by*

$$\frac{dC_t}{C_t} = \mu_c(\mathbf{s}_t)dt + \boldsymbol{\sigma}'_c \cdot d\mathbf{B}_t, \quad (5)$$

where

$$\mu_c(\mathbf{s}_t) = \bar{\mu}_c + \mathbf{s}'_t \cdot \boldsymbol{\theta} \quad (6)$$

where $\boldsymbol{\theta} = (\theta^1, \dots, \theta^n)'$, and $\boldsymbol{\sigma}_c = (\sigma_{c,1}, 0, \dots, 0)'$. The specification of θ^i is explained below.

Assumption 2. *The vector of consumption shares \mathbf{s}_t follows a continuous time, vector autoregressive process*

$$d\mathbf{s}_t = \boldsymbol{\Lambda}' \cdot \mathbf{s}_t dt + \mathbf{I}(\mathbf{s}_t) \cdot \boldsymbol{\Sigma}(\mathbf{s}_t) d\mathbf{B}_t, \quad (7)$$

where $\boldsymbol{\Lambda}$ is an $(n \times n)$ matrix with the property $\lambda_{ij} \geq 0$ for $i \neq j$, and $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$, and $\mathbf{I}(\mathbf{s}_t)$ is a diagonal matrix with i th element given by s_t^i , and $\boldsymbol{\Sigma}(\mathbf{s}_t)$ is an $(n \times n)$ matrix whose i th row is

$$\boldsymbol{\sigma}^i(\mathbf{s}_t) = \mathbf{v}'_i - \sum_{j=1}^n s_t^j \mathbf{v}'_j. \quad (8)$$

The cash-flow model, Equation (7), is extremely convenient while retaining a natural economic interpretation. The restrictions on the matrix $\boldsymbol{\Lambda}$ as well as the choice for the functional form of the volatility function, Equation (8), guarantee that both $s_t^i \geq 0$ and $\sum_{i=1}^n s_t^i = 1$ for all t . Thus, total income always equals consumption, dividends are never negative, and, for a generic matrix $\boldsymbol{\Lambda}$, no asset ever comes to fully fund consumption. In addition, fluctuations in the share of one asset may in turn induce variation in the fractions that other assets contribute to total consumption. Finally, cash-flow shocks naturally correlate across different assets, and the diffusion term in Equation (7) is flexible enough to capture a rich pattern in the covariance structure.⁶

⁶ For instance, if the assets are interpreted as industry portfolios, it is only natural that cash-flow shocks in a particular industry contain information about cash-flow growth in related industries.

The cash-flow model implied by Equation (7) generates an intuitive model for dividend growth. Indeed, an application of *Itô's lemma* shows that the process for dividends, \mathbf{D}_t , is given by

$$d\mathbf{D}_t = (\mathbf{I}(\bar{\mu}_c + \boldsymbol{\theta}) + \boldsymbol{\Lambda}') \cdot \mathbf{D}_t dt + \mathbf{I}(\mathbf{D}_t) \cdot \boldsymbol{\Sigma}_D(\mathbf{s}_t) d\mathbf{B}_t, \quad (9)$$

where $\mathbf{I}(\bar{\mu}_c + \boldsymbol{\theta})$ is a diagonal matrix with $\bar{\mu}_c + \theta^i$ in its i th position, the i th element of $\boldsymbol{\Sigma}_D(\mathbf{s}_t)$ is $\sigma_D^i(\mathbf{s}_t) = \boldsymbol{\sigma}_c + \boldsymbol{\sigma}^i(\mathbf{s}_t)$, and

$$\theta^i = \mathbf{v}'_i \cdot \boldsymbol{\sigma}_c.$$

That is, dividends follow a vector autoregressive process whose growth rates depend on the long-term growth of consumption itself, $\bar{\mu}_c$, as well as a parameter θ^i that regulates the instantaneous covariance between consumption growth and shares:

$$\text{cov}_t \left(\frac{ds_t^i}{s_t^i}, \frac{dC_t}{C_t} \right) = \theta^i - \sum_{j=1}^n s_t^j \theta^j \quad (10)$$

Thus, share processes that have a higher covariance with consumption growth will imply dividend streams with a higher growth in average. Notice that the constants θ^i are not identified as we can add a constant to all of them without changing any of the covariances. For this reason, we can renormalize them so that

$$\sum_{j=1}^n \theta^j \bar{s}^j = 0. \quad (11)$$

From Assumption 1, this condition implies that the unconditional expected consumption growth is $E[\mu_{c,t}] = \bar{\mu}_c + \sum_{j=1}^n \theta^j \bar{s}^j = \bar{\mu}_c$. In addition, this structure implies that the model is internally consistent. Applying the general formula for expected consumption growth, Equation (3), to Equation (9) we find

$$E_t \left[\frac{dC_t}{C_t} \right] = \sum_{i=1}^n s_t^i \mu_D^i(\mathbf{s}_t) = \sum_{i=1}^n s_t^i \frac{1}{D_t^i} ((\bar{\mu}_c + \theta^i) D_t^i + [\boldsymbol{\Lambda}' \cdot \mathbf{D}_t]_i) = \bar{\mu}_c + \sum_{i=1}^n s_t^i \theta^i$$

which equals Equation (6) in Assumption 1.⁷ We note that in the data, θ^i turn out to be very small. Thus, the conditional expected consumption growth is in fact essentially constant. In our simulations, we find that the minimum and maximum conditional expected consumption growth are just 2.25 and 2.39%, respectively, showing that the term $\sum_{i=1}^n s_t^i \theta^i$ has essentially no impact on $E_t[dC_t/C_t]$. On the other hand, this specification for expected consumption growth allows us to obtain closed form solutions for stock prices.

⁷ Here, we used the convenient fact that $\sum_{i=1}^n s_t^i / D_t^i [\boldsymbol{\Lambda}' \cdot \mathbf{D}_t]_i = 1/C_t \sum_{i=1}^n [\boldsymbol{\Lambda}' \cdot \mathbf{D}_t]_i = 0$, the latter equality stemming from the restrictions on the $\boldsymbol{\Lambda}$ matrix.

2. Results

2.1 Equilibrium prices

Given the consumption stream of the representative agent, the standard asset pricing formula is

$$P_t^i = E_t \left[\int_t^\infty \frac{U_c(\tau, C_\tau)}{U_c(t, C_t)} D_\tau^i d\tau \right]. \quad (12)$$

The appendix then proves the following proposition:

Proposition 1. *Under Assumptions 1 and 2, the price of asset i is given by*

$$P_t^i = \mathbf{b}'_i \cdot \mathbf{D}_t, \quad (13)$$

where \mathbf{b}'_i is the i th row of the matrix

$$\mathbf{b}' = (\mathbf{I}(\tilde{\phi}) - \Lambda')^{-1} \quad (14)$$

and $\mathbf{I}(\tilde{\phi})$ is the diagonal matrix with i th element given by

$$\tilde{\phi}^i = \phi - (1 - \gamma)\bar{\mu}_c + \frac{1}{2}\gamma(1 - \gamma)\boldsymbol{\sigma}'_c \boldsymbol{\sigma}_c - (1 - \gamma)\theta^i$$

This formula is very general. Because of the autoregressive nature of the cash-flow model, and the fact that consumption growth depends on the shares as well, as it should in a general equilibrium model [Equation (3)], the price of asset i does not depend only on its own dividend D_t^i , but also on the dividend level of all assets. Still, the effects of the asset's own dividends are first-order in the determination of its price. This can be seen from the dividend process itself [Equation (9)], as the i th element of the autoregressive model has the additional drift component $\bar{\mu}_c + \theta^i$, which determines the long-term properties of the dividend process itself. In addition, numerical examples show that the b_{ii} elements of the matrix \mathbf{b} are an order of magnitude larger than all the other entries.

Proposition 1 has an immediate implication for the value of the market and the total wealth portfolio, as they are simply $P_t^M = \sum_{i=2}^n P_t^i$ and $P_t^{TW} = \sum_{i=1}^n P_t^i$, respectively. In this case, we find that

$$P_t^M = \mathbf{b}'_M \cdot \mathbf{D}_t \quad \text{and} \quad P_t^{TW} = \mathbf{b}'_{TW} \cdot \mathbf{D}_t, \quad (15)$$

where $\mathbf{b}'_M = \sum_{i=2}^n \mathbf{b}'_i$ and $\mathbf{b}'_{TW} = \sum_{i=1}^n \mathbf{b}'_i$. From Equation (13) or (15), it is also immediate to compute the expected excess stock return. The excess stock returns of asset i , dR_t^i , are defined as

$$dR_t^i = \frac{dP_t^i + D_t^i dt}{P_t^i} - r_t dt.$$

The following result then applies to both individual assets and to the market portfolio:

Proposition 2. *The expected excess return on asset i is*

$$E_t[dR_t^i] = \gamma \left(\sigma_c' \sigma_c + \frac{\mathbf{b}_i' \cdot \mathbf{I}(\mathbf{s}_t)(\theta - 1_n \mathbf{s}_t' \cdot \theta)}{\mathbf{b}_i' \cdot \mathbf{s}_t} \right), \quad (16)$$

where $\mathbf{I}(\mathbf{s}_t)$ is the diagonal matrix with s_t^i in its i th entry.

In the case of the market portfolio, the formula holds substituting \mathbf{b}_M for \mathbf{b}_i throughout. Note that in a standard model with independently and identically distributed (i.i.d.) consumption and where consumption equals dividends, the equity premium would be given simply by $\gamma \sigma_c' \sigma_c$. The model proposed here is not only able to generate a time-varying equity premium as the shares move over time, but also a more sizable equity premium than that in a standard i.i.d. setting, thereby partly addressing the Mehra and Prescott (1985) equity premium puzzle.

2.2 An example

2.2.1 Assumptions and discussion. To sharpen intuitions about the role of labor income in asset pricing tests, it is useful to specialize the model further. First, we restrict the utility function to the log case $\gamma = 1$ as in the log CAPM of Rubinstein (1976). Second, we use a simple version of the cash-flow model, where the share of each asset depends only on its own past share value, and not on those of other assets. In addition, we assume that all financial assets are unconditionally identical, that is, all *financial* assets have the same covariance between share and consumption growth. The next assumption summarizes the essence of this simpler cash-flow model.

Assumption 3. (a) Let $\lambda_{ij} = a \bar{s}^j$ for $j \neq i$, where $\sum_{i=1}^n \bar{s}^i = 1$. Then the process governing s_t^i is given by

$$ds_t^i = a(\bar{s}^i - s_t^i)dt + s_t^i \boldsymbol{\sigma}^i(\mathbf{s}_t) \cdot d\mathbf{B}_t, \quad (17)$$

where $\boldsymbol{\sigma}^i(\mathbf{s}_t)$ is defined in Equation (8). (b) $\theta^i = \theta$ for $i = 2, \dots, n$.⁸

Assumption 3 has a natural interpretation for the cash-flow process: Asset i contributes a fraction s_t^i to overall consumption and, in the presence of shocks to s_t^i , it mean reverts to a long run value \bar{s}^i , which is the asset's steady state contribution to consumption. Second, the

⁸ Part (a) of Assumption 3 implies that $\lambda_{ii} = -\sum_{j \neq i} a \bar{s}^j = -a(1 - \bar{s}^i)$. This simpler model is the one used by Menzly, Santos, and Veronesi (2004), though they allow for cross-sectional variation in the covariance between share and consumption growth and in the speed of mean reversion, a . This case is also considered in Santos and Veronesi (2001).

relative share \bar{s}^i/s_t^i proxies for expected dividend growth. In fact, specializing Equation (9) to this case, we immediately obtain

$$E_t \left[\frac{dD_t^i}{D_t^i} \right] = \bar{\mu}_c + \theta^i + a \left(\frac{\bar{s}^i}{s_t^i} - 1 \right) \quad (18)$$

Third, financial assets have identical cash-flow risk, that is, the covariance of share and consumption growth is the same across these assets:

$$\text{cov}_t \left(\frac{ds_t^i}{s_t^i}, \frac{dC_t}{C_t} \right) = (\theta - \theta^w) s_t^w \quad \text{for } i = 2, \dots, n, \quad (19)$$

where $\theta^w = \theta^1$. This assumption means that, unconditionally, there will be no cross-sectional differences in prices–dividend ratios or in average returns. Still, as we show below, cross-sectional dispersion in expected dividend growth, as measured by \bar{s}^i/s_t^i , generates interesting conditional cross-sectional variation. Assumption 3 then is useful to investigate exactly how labor income interacts with dispersion in expected dividend growth to generate patterns in the cross-section of average returns and to understand the tests of the conditional CAPM where labor income is shown to play an important role. Finally, Assumption 3 together with Equation (10) implies that

$$\text{cov}_t \left(\frac{ds_t^w}{s_t^w}, \frac{dC_t}{C_t} \right) = -(\theta - \theta^w)(1 - s_t^w) \quad (20)$$

2.2.2 Price–dividend ratios. The next proposition follows immediately from Proposition 1.

Proposition 3. *The price of asset i is given by*

$$\frac{P_t^i}{D_t^i} = \left(\frac{1}{\phi} \right) \left(\frac{1}{\phi + a} \right) \left[\phi + a \left(\frac{\bar{s}^i}{s_t^i} \right) \right]. \quad (21)$$

Define next $P_t^M = \sum_{i=2}^n P_t^i$ and $D_t^M = \sum_{i=2}^n D_t^i$. Then

$$\frac{P_t^M}{D_t^M} = \left(\frac{1}{\phi} \right) \left(\frac{1}{\phi + a} \right) \left[\phi + a \left(\frac{1 - s_t^w}{1 - s_t^w} \right) \right] = \left(\frac{1}{\phi} \right) \psi(s_t^w). \quad (22)$$

Finally, the price of the total wealth portfolio, P_t^{TW} , is

$$\frac{P_t^{TW}}{C_t} = \frac{1}{\phi}. \quad (23)$$

Naturally, the price–dividend ratio of an asset i is an increasing function of the relative share. A high relative share \bar{s}^i/s_t^i implies a high

expected dividend growth [see Equation (18)] and thus the high price–dividend ratio.

As expression (22) shows, the result extends to the case of the market portfolio and it has a strong intuitive appeal. The first term of expression (22), $1/\phi$, is the “price–dividend” ratio of the total wealth portfolio [Equation (23)]. The second term, $\psi(s_t^w)$, corrects for the presence of an alternative source of income other than dividends from the market portfolio. Notice that $\psi(s_t^w) = 1$ only if $\bar{s}^w = s_t^w$. That is, an economy in its steady state yields a price–dividend ratio that is no different than the usual one. Deviations from this steady state generate movements in the price–dividend ratio of the market portfolio. For instance, if $\bar{s}^w < s_t^w$ then the price–dividend ratio is higher than its long run level, $1/\phi$. There are two reasons for this. First, if s_t^w is relatively high, investors are less exposed to fluctuations in the stock market, and hence, they require a lower compensation to hold it; this, in turn, translates into higher prices. Second, a high share of labor income to consumption signals that future aggregate dividend growth is going to be above that of consumption as s_t^w will mean revert to \bar{s}^w . This further reinforces the positive effect on the price–dividend ratio.

2.2.3 Expected excess returns.

Proposition 4. *The expected excess returns of asset i and the market portfolio are respectively given by*

$$E_t[dR_t^i] = \sigma_c' \sigma_c + \left[\frac{\theta - \theta^w}{1 + \frac{a}{\phi} \left(\frac{\bar{s}^i}{s_t^i} \right)} \right] s_t^w \quad (24)$$

and

$$E_t[dR_t^M] = \sigma_c' \sigma_c + (\theta - \theta^w) \left(\frac{s_t^w (1 - s_t^w)}{\phi(1 - s_t^w) + a(1 - \bar{s}^w)} \right) \quad (25)$$

To understand Equation (24), recall first that in the log economy the expected excess return of the total wealth portfolio, $E_t[dR_t^{TW}]$, is given by $\sigma_c' \sigma_c$. The expected excess return of asset i is above or below $E_t[dR_t^{TW}]$ depending on whether the covariance between share and consumption growth, $-(\theta - \theta^w)s_t^w$, is positive or negative [Equation (19)]. Evidence reported below show that the correlation between consumption and share growth is negative and hence $(\theta - \theta^w) > 0$. It follows then that under Assumption 3, $\text{cov}(ds_t^i/s_t^i, dC_t/C_t) > 0$ and that an asset’s contribution to consumption grows precisely when consumption does, which makes the asset risky. For this reason, financial assets command a premium over that of the total wealth portfolio, $\sigma_c' \sigma_c$. The expected excess

return of asset i is also determined by the expected dividend growth as proxied by the relative share \bar{s}^i/s_t^i [see Equation (18)]. An asset with a large current share commands a higher premium than an otherwise identical asset with a lower share, as it is a larger fraction of consumption and thus riskier.

The degree to which changes in s_t^w affect $E_t[dR_t^i]$ depends also on the value of s_t^i . If $s_t^i \approx 0$, changes in s_t^w do not affect the required return, as asset i does not contribute to consumption and hence does not co-vary with its growth. Notice that \bar{s}^i/s_t^i is high for stocks that pay in the future, which could be termed “growth” stocks. These assets then will have both low expected returns and a relatively lower sensitivity to changes in s_t^w .

Equation (24) shows that, in the context of the present model, whatever cross-sectional dispersion in average returns that is observed empirically, can only spring from the *conditional* dispersion in expected excess returns. The model generates this conditional cross-sectional dispersion through temporary shocks to expected dividend growth. The role of labor income as a useful variable in tests of the conditional CAPM can now be easily illustrated. For instance, sorting portfolios by the price–dividend ratio is, in this model, equivalent to sorting by the relative share, \bar{s}^i/s_t^i [see Equation (21)]. This sorting captures the source of (conditional) dispersion in the cross-section of returns. Still the size of this dispersion is governed by the share of labor income to consumption, s_t^w . If s_t^w is very small, whatever dispersion there is in price–dividend ratios, does not translate into a large cross-sectional dispersion of returns. Conversely, a small dispersion in price–dividend ratios can translate into a large dispersion of returns when s_t^w is large. It follows that the inclusion of labor income, normalized by consumption in this case, in tests of the cross-section can help align portfolios, particularly when these are sorted according to some valuation ratio.

As for the market portfolio equation, Equation (25) shows that the instantaneous expected return depends nonlinearly on the fraction of consumption produced by labor income $s_t^w = w_t/C_t$. Its functional form shows that expected returns are equal to $\sigma_c' \sigma_c$ both when $s_t^w = 0$ and when $s_t^w = 1$. Indeed, when $s_t^w = 0$, then $C_t = \sum_{j=2}^n D_t^j = D_t^M$ and we revert to an economy with no other endowment than the risky assets. In this case $E_t[dR_t^M] = \sigma_c' \sigma_c$, the standard equity premium in the log economy. More puzzling perhaps at first is that this model implies that $E_t[dR_t^M] = \sigma_c' \sigma_c$ when $s_t^w = 1$. In this case we must have $s_t^i = 0$ for all $i = 2, \dots, n$ and thus

$$P_t^M = C_t \frac{a}{\phi(a + \phi)} (1 - \bar{s}^w).$$

Since, in this case, we also have that $C_t = w_t$, the price is perfectly correlated with wages (and hence consumption), yielding the result.

Clearly, these two cases are extreme, given that, as shown below s_t^w lies comfortably in the interval (0.7, 0.95) in the postwar sample. In this case,

what is the relationship between $E_t[dR_t^M]$ and s_t^w ? In Equation (25) the denominator of the second term is always positive, hence the behavior of expected stock returns depends solely on the sign of $\theta - \theta^w$, which is positive, as discussed earlier. This is also economically intuitive: If wages are much smoother than dividends, an increase in dividends is accompanied by an increase in consumption and hence a decrease in $s_t^w = w_t/C_t$. This induces a negative covariance between consumption growth and changes in s_t^w which, from Equation (20), implies $\theta - \theta^w > 0$. This yields in turn a negative relation between expected returns and the labor share s_t^w when s_t^w is in the relevant range (0.7, 0.95). The economic intuition of this result is also clear: As s_t^w increases, consumption becomes fueled by labor income only, decreasing the covariance between consumption growth and dividend growth. This, in turn, translates into a lower covariance between consumption growth and returns, generating a lower-risk premium. Thus, a high labor income to consumption ratio should forecast low future excess returns.

2.2.4 The CAPM representation. The log utility case also has the advantage of providing simple and intuitive formulas for the CAPM representation of expected returns. As shown in Proposition 3, $P_t^{TW} = \phi^{-1}C_t$. This implies that this asset is perfectly correlated with the stochastic discount factor and thus, from standard results [e.g., Duffie (1996)], a CAPM representation holds. Specifically, we have

Proposition 5. *Expected excess returns on individual securities are given by*

$$E_t[dR_t^i] = \beta^{TW} \left(\frac{\bar{s}^i}{s_t^i}, s_t^w \right) E_t[dR_t^{TW}], \quad (26)$$

where

$$\beta^{TW} \left(\frac{\bar{s}^i}{s_t^i}, s_t^w \right) = \frac{\text{cov}_t(dR_t^i, dR_t^{TW})}{\text{var}_t(dR_t^{TW})} = 1 + \left(\frac{\theta - \theta^w}{\sigma_c' \sigma_c \left(1 + \frac{\alpha}{\phi} \left(\frac{\bar{s}^i}{s_t^i} \right) \right)} \right) s_t^w. \quad (27)$$

Equation (26) is the CAPM with respect to the total wealth portfolio. β^{TW} depends on both a common factor, which is the labor income to consumption ratio, and an asset's characteristic, the relative share. Thus, the presence of nonfinancial sources of income induces time variation in the asset's beta. The intuition, of course, for Equation (27) is identical to the one above for Equation (24) and we omit it in the interest of space. Tests of Equation (26) require observation of the total wealth portfolio, which is difficult. The appendix shows the following:

Proposition 6. *The expected excess returns on individual securities are given by*

$$E_t[dR_t^i] = \beta^{w,i}(\mathbf{s}_t)E_t[dR_t^w] + \beta^{M,i}(\mathbf{s}_t)E_t[dR_t^M], \quad (28)$$

where $\beta^{w,i}(\mathbf{s}_t)$ and $\beta^{M,i}(\mathbf{s}_t)$ are the multiple regression coefficients,

$$[\beta^{w,i}(\mathbf{s}_t) \ \beta^{M,i}(\mathbf{s}_t)]' = (\boldsymbol{\Sigma}^{wM})^{-1} \times [\text{cov}_t(dR_t^i, dR_t^w) \ \text{cov}_t(dR_t^i, dR_t^M)]'$$

and where $\boldsymbol{\Sigma}^{wM}$ is the variance–covariance matrix of dR_t^w and dR_t^M .

Versions of Equation (28) have been the focus of much research lately. For instance, Jagannathan and Wang (1996) test a version of the above equation, where they also extend the definition of the market portfolio to include returns to human capital, and where their conditioning variable is the properly defined default premium, shown to forecast business cycles. More recently, Lettau and Ludvigson (2001b) have tested a similar equation in a different set of test portfolios where the conditioning variable is the consumption to wealth ratio, a conditioning variable that they show predicts future market returns.

2.3 Predictability when $\gamma > 1$

The intuition developed in Section 2.2 carries through to the case where $\gamma > 1$. In what concerns the time-series predictability, it is easy to see from Equation (16) that under Assumption 3 the expected excess return of any asset is only a function of s_t^i and s_t^w . This immediately implies that, as in the log utility case, the expected excess return on the market portfolio is only a function of s_t^w and thus this latter variable should forecast future excess returns.

As for the cross-section of stock returns, and as was the case with log utility, sorting by the price–dividend ratio is akin to sorting by expected excess returns. In fact, it is possible to show that when $\gamma > 1$ a stock with a high expected dividend growth is still characterized by a high price–dividend ratio and a low expected excess return. As in the log utility case, then, the sorting procedure effectively removes differences in conditional expected excess returns that are due to cross-sectional variation in \bar{s}^i/s_t^i . Thus, the remaining variation in the cross-sectional dispersion of returns is only driven by s_t^w , which acts as a common factor across stocks.

Unlike the log utility case though, when $\gamma > 1$ the conditional CAPM with respect to the total wealth portfolio does not hold. The reason is that the correlation between consumption shocks and the returns on the total wealth portfolio is not one. Still, the simulation exercise performed in Section 3.3.4 shows that this correlation is as high as 90%, and thus, little is lost by assuming that the conditional CAPM holds *approximately* when $\gamma > 1$. Indeed, in line with this latter finding and the intuition above, we

also show in Section 3.3.4 that a version of the conditional CAPM that uses the share of the labor income to consumption as a conditioning variable can explain essentially all of the cross-sectional variation in average returns, whereas the unconditional CAPM leaves much of this variation unexplained.

3. Empirical Results

3.1 Data description

We consider returns on the value-weighted CRSP index, which includes NYSE, AMEX, and NASDAQ, as our measure of financial asset returns. Dividend–price ratios are also obtained from CRSP and the risk-free rate is the 90-day Treasury bill. For the cross-sectional tests, we use the set of portfolios constructed by Fama and French (1993), formed by intersecting five size-sorted portfolios with five other portfolios sorted by book-to-market.

As for the macroeconomic time series they are all obtained from the Bureau of Economic Analysis. Consumption (C_t) is defined as nondurables plus services. We use the definition of labor income (w_t) in Lettau and Ludvigson (2001a), which is, briefly, wages and salaries plus transfer payments plus other labor income minus personal contributions for social insurance minus taxes. Both the consumption and labor income series are quarterly and our sample period is 1948–2001. With these two series we construct our main state variable, the share of labor income to consumption $s_t^w = w_t/C_t$. As a robustness check, we also use two alternative constructions of the share of labor income to consumption. First, we use compensation of employees (w_t^{ce}) as our definition of labor income, which is computed by adding wage and salary accruals plus supplements to wages and salaries (employer contributions for social insurance plus other labor income.) The labor income to consumption ratio is then computed as the ratio of the compensation of employees to the consumption of nondurables plus services, $s_t^w = w_t^{ce}/C_t$. Second, general equilibrium models blur the distinction between consumption and income. Thus, a normalization of labor income by disposable income, w_t/Y_t , is also theoretically sound and we employ it below as well. Figure 1 plots these three series for our sample period.

Table 1 provides some summary statistics. As can be seen, our measures of s_t^w are all highly persistent and thus caution has to be exercised when drawing inferences about the forecasting ability of these variables. Also, as discussed in Section 2.2, the impact of changes in the share of labor income to consumption on the expected excess return of the market portfolio depends critically on the sign of the correlation between ds_t^w and dC_t . As can be seen in panel B, independent of the particular measure of the share of labor income to consumption employed, this correlation is

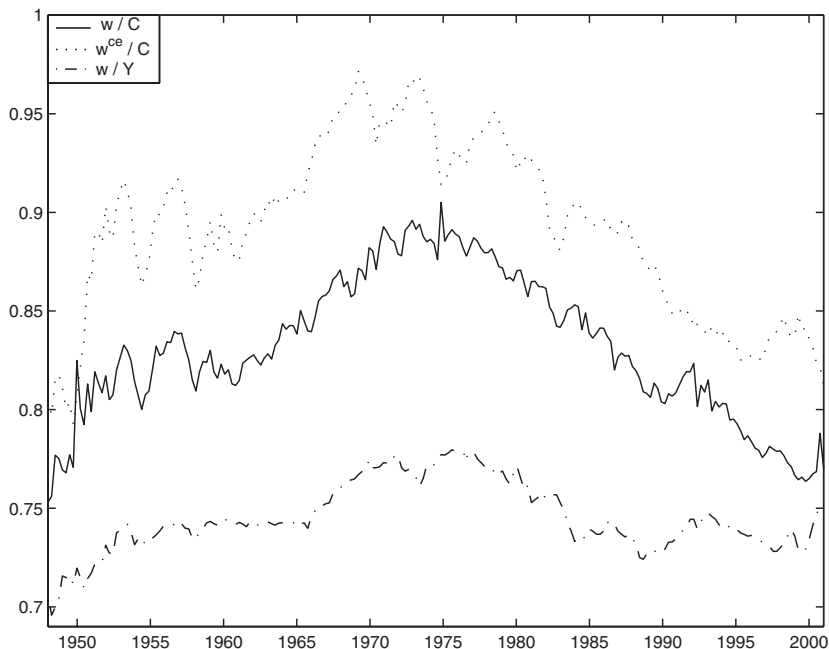


Figure 1
Shares of labor income to consumption

Current share of labor income to consumption, s_t^w , defined as (i) the ratio of labor income over consumption w/C (solid line), (ii) the ratio of compensation of employees over consumption w^{ce}/C (dotted line), or (iii) the ratio of labor income over disposable income w/Y (dash-dotted line). Data is quarterly and the sample period is 1948–2001.

negative. Thus, for the purposes of the interpretation of the empirical results below in light of our findings in Section 2.2, $\theta - \theta^w > 0$.

Panel C reports results of an Augmented Dickey Fuller (ADF) test for the presence of a unit autoregressive root in $\log(s_t^w) = \log(w_t) - \log(C_t)$, which our model assumes is stationary. The results are reported for different choices regarding the lags ℓ in

$$\log(s_t^w) = a_0 + a_1 \log(s_{t-1}^w) + \sum_{j=0}^{\ell} \zeta_j \Delta \log(s_{t-j}^w) + \varepsilon_t$$

[see Hamilton (1994), chap. 17] The optimal number of lags is chosen according to a sequential procedure described in Campbell and Perron (1991) and in Ng and Perron (1995), and it is denoted by an asterisk in the appropriate entry in the panel. For the sample 1948–2001, we cannot reject the hypothesis that $\log(s_t^w)$ follows a unit root process at the standard confidence levels. Since a reasonable concern lies in the low power of the ADF test in relatively small samples, it is useful to ascertain

Table 1
Basic statistics

	$\Delta \log(C)$	R_t^M	$\ln(D/P)$	w_t/C_t	w_t^{ce}/C_t	w_t/Y_t	
Panel A: Summary statistics							
Mean	0.0056	0.0195	-3.3783	0.8317	0.8918	0.7454	—
Standard deviation	0.0054	0.0814	0.3937	0.0362	0.0438	0.0174	—
Autocorrelation coefficient	0.2131	0.0432	0.9869	0.9755	0.9857	0.9755	—
β	0.2131	0.0433	0.9946	0.9712	0.9842	0.9723	—
	$\Delta \log(C)$	R_t^M	$\Delta \ln(D/P)$	$\Delta(w_t/C_t)$	$\Delta(w_t^{ce}/C_t)$	$\Delta(w_t/Y_t)$	
Panel B: Correlation matrix							
$\Delta \log(C)$	1	0.2054	-0.2454	-0.1724	-0.0761	-0.1466	—
R_t^M	0.2054	1	-0.2626	0.0719	0.1155	-0.0026	—
$\Delta \ln(D/P)$	-0.2454	-0.2626	1	0.0160	0.0502	0.0659	—
$\Delta(w_t/C_t)$	-0.1724	0.0719	0.0160	1	0.4680	0.3337	—
$\Delta(w_t^{ce}/C_t)$	-0.0761	0.1155	0.0502	0.4680	1	0.1700	—
$\Delta(w_t/Y_t)$	-0.1466	-0.0026	0.0659	0.3337	0.1700	1	—
Lags	0	1	2	3	4	5	6
Panel C: Augmented Dickey-Fuller test for Unit Root: 1948–2001							
w_t/C_t	-1.968	-1.683*	-1.198	-1.168	-1.141	-1.211	-1.033
w_t^{ce}/C_t	-1.390	-1.848	-1.603	-1.651	-1.611*	-1.904	-1.701
w_t/Y_t	-2.325*	-3.081	-2.757	-2.576	-1.967	-2.087	-2.069
Lags	0	1	2	3	4	5	6
Panel D: Augmented Dickey-Fuller test for Unit Root: 1929–2001							
w_t/C_t	-2.180	-2.202	-2.388	-2.721*	-2.605	-2.684	-2.823
w_t^{ce}/C_t	-2.187	-2.300	-2.633*	-2.758	-2.691	-2.819	-2.948
w_t/Y_t	-4.028	-3.244*	-3.101	-2.985	-3.096	-3.057	-2.763

Panel A: summary statistics for $\Delta \log(C)$, aggregate consumption growth, R^M , market returns, $\ln(D/P)$, the log of the dividend yield, and s_t^w , the labor income to consumption ratio. This is defined as labor income, as defined in Lettau and Ludvigson (2001a) w_t , divided by consumption of nondurables and services C_t , “compensation of employees” w_t^{ce} divided by consumption, and as labor income divided by disposable income Y_t . The last line reports the value of the regression coefficient, β , of an OLS regression on own lagged variable. Panel B: correlation matrix, where a Δ in front a variable denotes the first difference operator. Panels C and D: results of the Augmented Dickey Fuller (ADF) test for a unit autoregressive root in $\log(s_t^w) = \log(w_t) - \log(C_t)$ for the sample 1948–2001 and 1929–2001, respectively and for each of the alternative definitions of s_t^w . “Lags” refers to the number of lags ℓ in the regression $\log(s_t^w) = a_0 + a_1 \log(s_{t-1}^w) + \sum_{j=0}^{\ell} \zeta_j \Delta \log(s_{t-j}^w) + \varepsilon_t$. The optimal number of lags, denoted by superscript *, is chosen by using the sequential step-down procedure described in Ng and Perron (1995). Critical values at 1, 5, and 10% are given by -3.4926, -2.8760, and -2.5688 for the 1948–2001 sample, and by -3.4404, -2.8697, and -2.5829, for the 1929–2001 sample. Data and units are quarterly and the sample period is 1948–2001, except in panel D where the sample is 1929–2001.

whether the same result is obtained when we use a longer sample period. To this end, we interpolate the annual data in the NIPA tables that are now available for the period 1929–1948 to obtain a quarterly series of s_t^w for this longer sample period. The results are contained in panel D. In this case we can reject the null of a unit root at the 10% level for the case where s_t^w is built using either the labor income measure of Lettau and Ludvigson (2001a) or the compensation of employees. Instead when s_t^w is measured as the labor income to disposable income ratio, we can reject a unit root at the 5% level. Although the evidence is mixed, the power of these tests is,

as mentioned, low and the restriction that $\log(s_t^w)$ is stationary rests on solid economic intuition: it is not reasonable to assume that consumption can grow to be infinitely larger than labor income, or, alternatively, that labor income can grow to be several times higher than consumption. Our model Equations (5)–(8) is simply a tractable way of capturing this basic economic intuition.

3.2 Predictability of aggregate returns

3.2.1 s_t^w and the predictability of long-horizon returns. The main prediction of our model is that a high share of labor income to consumption, s_t^w , predicts low future returns. Panels A and B of Figure 2 give a visual impression of the behavior of both the share of labor income to consumption and the log dividend yield versus long-horizon returns, measured by the four-year cumulative returns. The share of labor income to consumption ratio does indeed move in the opposite direction to the long-horizon returns, as predicted by theory. As for the log dividend yield and long-horizon returns, the first part of the sample shows the familiar

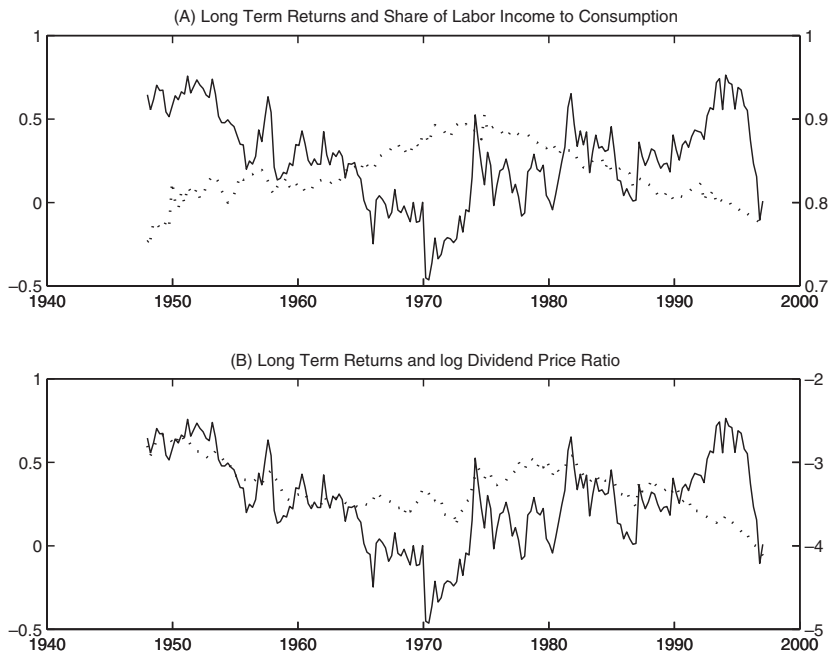


Figure 2

Long-term returns and predictive variables

Four-year cumulative market returns (solid line) lagged four years and the current share of labor income to consumption (dotted line, panel A), and the log dividend–price ratio (dotted line, panel B). Data is quarterly and the sample period is 1948–2001.

pattern in the predictability literature: log dividend yields co-move with long-horizon returns. The plot also shows the striking behavior of the market during the 1990s. Contrary to the prior historical experience, log dividend-price ratios and long-horizon returns move in opposite direction during that period until the correction of 2001.

In panel A of Table 2, we report the results of regressions of long-horizon excess returns on lagged values of s_t^w on the log dividend yield and on both. That is, we run

$$\text{Regression 1 } r_{t,t+K} = \alpha_1 + \beta_1(K)s_t^w + \varepsilon_{t+k} \quad (29)$$

$$\text{Regression 2 } r_{t,t+K} = \alpha_2 + \beta_2(K)\log\left(\frac{D_t^M}{P_t^M}\right) + \varepsilon_{t+k} \quad (30)$$

$$\text{Regression 3 } r_{t,t+K} = \alpha_3 + \beta_3(K)s_t^w + \beta_4(K)\log\left(\frac{D_t^M}{P_t^M}\right) + \varepsilon_{t+k}, \quad (31)$$

Table 2
Forecasting regressions— $s_t^w = w_t/C_t$

	Panel A: Sample 1948–2001				Panel B: Sample 1948–1994			
Regression 1								
Horizon	4	8	12	16	4	8	12	16
s_t^w	-0.93 (-1.44)	-2.48* (-3.07)	-4.01* (-4.24)	-5.25* (-4.92)	-1.39* (-2.30)	-3.04* (-4.33)	-4.41* (-4.39)	-5.54* (-4.72)
	[-1.38]	[-1.93]	[-2.09]	[-2.14]	[-1.89]	[-2.11]	[-2.07]	[-1.99]
Adjusted R^2	0.04	0.16	0.32	0.42	0.07	0.21	0.35	0.44
Regression 2								
Horizon	4	8	12	16	4	8	12	16
$\log(D/P)$	0.13* (2.13)	0.20 (1.65)	0.26 (1.34)	0.35 (1.29)	0.28* (4.04)	0.48* (4.00)	0.63* (4.49)	0.78* (5.41)
	[2.24]	[1.75]	[1.49]	[1.60]	[3.66]	[2.95]	[2.50]	[2.43]
Adjusted R^2	0.09	0.10	0.11	0.14	0.19	0.32	0.43	0.54
Regression 3								
Horizon	4	8	12	16	4	8	12	16
s_t^w	-1.43* (-2.49)	-2.97* (-4.56)	-4.31* (-5.56)	-5.30* (-6.06)	-0.83 (-1.51)	-2.00* (-3.72)	-3.03* (-4.92)	-3.72* (-5.34)
	[-2.20]	[-2.32]	[-2.25]	[-2.17]	[-1.08]	[-1.36]	[-1.40]	[-1.27]
$\log(D/P)$	0.17* (2.96)	0.26* (3.31)	0.30* (3.54)	0.35* (3.25)	0.25* (4.12)	0.40* (3.90)	0.49* (4.99)	0.60* (7.15)
	[2.97]	[2.40]	[1.85]	[1.75]	[3.12]	[2.37]	[1.88]	[1.75]
Adjusted R^2	0.17	0.33	0.47	0.57	0.21	0.40	0.57	0.71

The table summarizes the result of the predictive regression

$$r_{t,t+K} = \alpha + \beta(K)\mathbf{x}_t + \varepsilon_{t+K},$$

where $x_t = s_t^w$, $\log(D_t/P_t)$, or both, where K is the number of quarters ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters and $s_t^w = w_t/C_t$. Here, w_t is labor income as defined in Lettau and Ludvigson (2001a) and C_t is aggregate consumption of nondurable goods and services. Numbers in parenthesis show the Newey–West adjusted t -statistics, where the number of lags is double that of the forecasting horizon. The t -statistics computed with the Hodrick (1992) type 1B standard errors are reported in brackets. Data are quarterly and the sample is 1948–2001 (panel A) and 1948–1994 (panel B). *significance at the 5% level using the Newey–West adjusted t -statistic.

where $r_{t,t+K}$ is the cumulative log excess return over K periods. For each regression, we report the point estimates of the included explanatory variable, the adjusted R^2 s and, in parentheses, the Newey–West corrected t -statistic. As already mentioned, the share of labor income is highly persistent and this complicates the inference on the significance of the estimated coefficient. For this reason we also report, in brackets, the t -statistic obtained using the standard errors proposed by Hodrick (1992).⁹ Ang and Bekaert (2001) show that these standard errors have better small sample properties than alternative ones and thus we compute them throughout.

Turning first to the regression of long-horizon returns on s_t^w , panel A of Table 2 shows that this variable is a statistically significant predictor of returns for horizons of two years and longer when the Newey–West t -statistic is considered. The significance remains at horizons of three and four years when using the Hodrick t -statistics, being only marginally so for the two-year horizon. Consistent with the theory and the negative correlation between changes in shares and consumption growth, the sign of the coefficient is negative: positive innovations in s_t^w lead to low future returns. The explanatory power is also high, ranging from 16% for the two-year regression to 42% for the four-year regression, displaying the familiar increasing pattern with respect to the forecasting horizon.

The dividend yield does not forecast returns as well and it is significant only at the one-year horizon, precisely where s_t^w was not. When both regressors are included, as in regression (31), the share of labor income to consumption becomes significant at all horizons independently of whether one considers the Newey–West t -statistic or the more stringent Hodrick t -statistic. The significance of the log dividend yield improves instead only marginally. Finally, the R^2 s increase considerably over the case where only the share of labor income to consumption is included and reaches a remarkable 57% at the four-year horizon.

3.2.2 Robustness to alternative definitions of s_t^w and sample periods. The share of labor income to consumption forecasts returns at long horizons, whereas the dividend yield, contrary to received wisdom from the predictability literature, does much more poorly. Indeed, as already mentioned and strikingly documented in Figure 2, the late 1990s saw high returns and low dividend yields, and this has seriously diminished the ability of the dividend yield to forecast long-horizon returns. The question then remains whether the predictability of the share of labor income is somehow related to its performance during this peculiar period. To address this question, panel B of Table 2 reports the results of regressions

⁹ Hodrick (1992) refers to the standard errors he proposes as “standard errors (IB)” (page 362.) We follow Ang and Bekaert (2001) in referring to them as Hodrick standard errors.

(29)–(31), but now run in a shorter sample that excludes the extraordinary market of the second half of the 1990s (1948–1994).

The log dividend yield is now a strongly significant predictor of stock returns, independently of whether one draws the inference considering the Newey–West or the Hodrick *t*-statistics. The lack of predictability of the dividend yield in the complete sample is an artifact of the stock market performance during the late 1990s. Notice as well that the R^2 has increased considerably to the levels that are usual in the predictability literature. The significance of s_t^w as a strong predictor of long-horizon returns remains in this shorter sample and is robust to the inclusion of the dividend yield, though it is slightly lower if the Hodrick *t*-statistics are considered to draw inferences.

As for the construction of s_t^w , our definition of labor income and consumption is standard [see, e.g., Lettau and Ludvigson (2001a)], but clearly, given that theory is silent on the specifics of this construction, it is useful to know whether its forecasting ability is robust to the two alternative measures of s_t^w introduced in Section 3.1.

First, panel A of Table 3 uses $s_t^w = w_t^{ce}/C_t$ as our definition of the share of labor income to consumption, where recall w_t^{ce} is defined as compensation of employees, rather than labor income. The results are very similar

Table 3
Forecasting regressions—robustness

	Panel A: $s_t^w = w_t^{ce}/C_t$				Panel B: $s_t^w = w_t/Y_t$			
Regression 1								
Horizon	4	8	12	16	4	8	12	16
s_t^w	-1.04*	-2.13*	-3.22*	-4.22*	-2.27*	-5.23*	-7.38*	-9.50*
	(-2.38)	(-4.17)	(-6.02)	(-6.16)	(-2.34)	(-4.50)	(-4.43)	(-4.59)
	[-1.96]	[-2.08]	[-2.18]	[-2.33]	[-1.85]	[-2.23]	[-2.12]	[-2.06]
Adjusted R^2	0.07	0.18	0.32	0.44	0.06	0.19	0.29	0.39
Regression 2								
Horizon	4	8	12	16	4	8	12	16
s_t^w	-1.32*	-2.41*	-3.40*	-4.24*	-2.17*	-4.90*	-6.88*	-8.68*
	(-2.97)	(-4.59)	(-5.32)	(-5.39)	(-2.29)	(-4.26)	(-4.87)	(-5.21)
	[-2.52]	[-2.35]	[-2.30]	[-2.35]	[-1.78]	[-2.08]	[-1.97]	[-1.88]
log(D/P)	0.16*	0.24*	0.29*	0.35*	0.13*	0.18	0.20	0.24
	(2.97)	(3.25)	(3.70)	(3.88)	(2.21)	(1.92)	(1.67)	(1.39)
	[2.83]	[2.26]	[1.80]	[1.76]	[2.16]	[1.57]	[1.20]	[1.11]
Adjusted R^2	0.20	0.33	0.47	0.58	0.14	0.26	0.36	0.45

The table summarizes the result of the predictive regression

$$r_{t,t+K} = \alpha + \beta(K)\mathbf{x}_t + \varepsilon_{t+K},$$

where $x_t = s_t^w$, or both s_t^w and $\log(D/P_t)$, and where K is the numbers of quarters ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. s_t^w is defined either as compensation to employees w_t^{ce} divided by consumption of nondurable goods and services C_t (panel A) or as labor income w_t divided by disposable income Y_t (panel B). Numbers in parenthesis show the Newey–West adjusted *t*-statistics, where the number of lags is double that of the forecasting horizon. The *t*-statistics computed with the Hodrick (1992) type 1B standard errors are reported in brackets. Data are quarterly and the sample is 1948–2001. *significance at the 5% level using the Newey–West adjusted *t*-statistic.

to the previous ones if not stronger. s_t^w is statistically significant at the 5% level at all horizons, independent of the standard errors used to compute the t -statistics. The R^2 s are of very similar magnitudes to those in Table 1. Once again, the significance of s_t^w is there even when controlling for the log dividend yield. We also run regressions 1 and 3 using $s_t^w = w_t/Y_t$, that is, normalizing labor income by disposable income. As shown in panel B of Table 3, the significance of s_t^w is unaffected, except when one considers the one-year horizon and computes the t -statistic with the Hodrick standard errors, where the significance is only at the 10% level. These results extend to the case where both s_t^w and $\log(D_t^M/P_t^M)$ are included though now the significance of the log dividend–price ratio is weaker if the Hodrick standard errors are used.¹⁰

3.2.3 Simulation results. *Details of the simulation procedure.* In this section, we report results from simulating the model described in Section 1. Specifically, we generate 10,000 years of artificial quarterly data for the consumption growth process as well as for the shares of 200 identical assets.¹¹ We use these fictitious assets to construct both the market portfolio and a set of price-sorted portfolios to test, in artificial data, both the time-series and cross-sectional implications of our model and compare them to their empirical counterparts. Panel A of Figure 3 shows the histogram of the generated process for the share of labor income to consumption ratio. Notice that the bulk of the mass concentrates in the historical range and, as the thin left-hand side tail shows, s_t^w wanders as low as 0.5 only very rarely.

Panel A of Table 4 contains the parameters for the consumption growth and share processes, Equations (5) and (17), respectively. It also reports our choices for the preference parameters, γ and ϕ . We set $\gamma = 60$ for, as it is well known, a high level of risk aversion is needed to match the equity premium in consumption-based models even in those that depart

¹⁰ As discussed in Santos and Veronesi (2001), we performed additional robustness checks. First, we used Monte Carlo simulations to compute standard errors that are robust to spurious regression. We found that s_t^w is significantly negative at the 5% level in all samples and for all constructions. Second, we also tested directly whether the covariance between consumption growth and returns is negatively related to s_t^w . We used the Vech–Garch model of Bollerslev, Engle, and Woodridge (1988) to compute the time series of this conditional covariance, and then regress it on s_t^w . We found a negative slope coefficient, as predicted by our model. The coefficient though is not significant if the consumption deflator is used to obtain real consumption, whereas it is significant if the CPI is used. Using different data and monthly frequency, Duffee (2004) finds instead an insignificant positive relation. We conclude that direct evidence on the conditional covariance between return and consumption growth is weak, a result that is to be expected given the noise in the real consumption growth data series (especially at the monthly frequency).

¹¹ The fact that all assets are *ex ante* identical implies that only the properties of the wages–consumption ratio s_t^w matter for the market portfolio and not the number or the characteristics of individual securities. To see this, notice that for any two assets $i, j \geq 2$, we have $b_i^j = b_j^i$ and $b_i^k = b_j^k$ for $k \neq i, j$. Thus, it is easy to see that we can write $P_t^M = \sum_{i=2}^n P_t^i = C_i(b_M + (b_M^1 - b_M)s_t^w)$ for two constants b_M and b_M^1 .

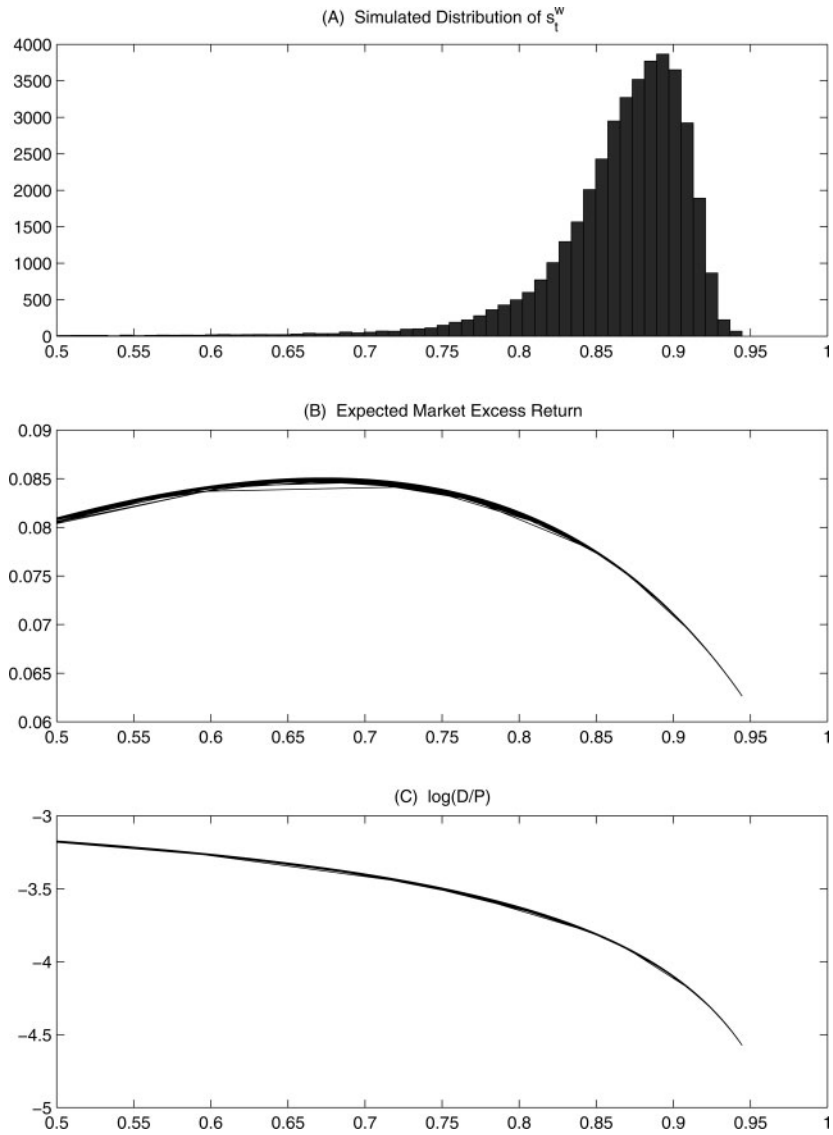


Figure 3
Simulations

Panel A: histogram of simulated share $s_t^w = w_t/C_t$ (see notes to Table 4 for details on simulation). Panel B: theoretical expected return plotted against the labor income share s_t^w . Panel C: model implied log dividend–price ratio as a function of labor income share s_t^w .

Table 4
Simulation parameters

Panel A: Parameters										
γ	ϕ	μ_c	σ_c	\bar{s}^w	\bar{s}^i	a	$\nu^{w,1}$	$\nu^{w,2}$	$\nu^{j,1}$	$\nu^{j,i+1}$
60	0.0984	0.0223	0.0280	0.8637	0.0007	0.0892	-0.0089	-0.0826	0.0567	0.5934
Panel B: Unconditional moments from simulations										
$E(s_t^w)$	Std(s_t^w)	Corr(ds_t^w, dc_t)	$E(R_t^M)$	Std(R_t^M)	$E(r_t^f)$	Std(r_t^f)				
0.8637	0.0556	-0.2412	0.0753	0.0621	0.03	0.0007	—	—	—	—

Parameters used for the simulation of the consumption and share processes. We simulate 10,000 years of quarterly data of consumption, labor share, and shares for 200 identical assets. Panel A: parameters used in simulations, where γ , coefficient of relative risk aversion; ϕ , subjective discount rate; μ_c , annualized average consumption growth; and σ_c , annualized standard deviation of consumption growth. The parameters used for the share processes are as follows. \bar{s}^w and \bar{s}^i , long run contribution of both labor income and asset i to consumption; a , speed of mean reversion of the share process; $\nu^{w,1}$ and $\nu^{w,2}$, first and second entries of the vector ν^w and $\nu^{w,j} = 0$ for $i = 3, \dots, n$; $\nu^{j,1}$ and $\nu^{j,i+1}$ for $i = 2, \dots, n$ denote the first and $i + 1$ th entry of the vector ν^j , and $\nu^{j,j+1} = 0$ for $j \neq 1, i$. Panel B: unconditional moments obtained using simulated data. $E(\bar{s}^w)$, average contribution of labor income to consumption, $\text{std}(s_t^w)$, standard deviation of the share of labor income to consumption; $\text{corr}(ds_t^w, dc_t)$, correlation between changes in the share s_t^w and log consumption $c_t = \log(C_t)$; $E(R^M)$, average excess return on the market portfolio, whereas $\text{std}(R^M)$, standard deviation; $E(r^f)$ and $\text{std}(r^f)$ are the mean and the standard deviation of the risk-free rate. The time step used in simulations is $dt = 0.0208$.

from the traditional CRRA preferences. For instance, the habit formation model of Campbell and Cochrane (1999), a model that is able to match many features of the data, implies a risk aversion level around 80. A relative risk aversion equal to 60 may appear too high—even higher than the one used by Mehra and Prescott (1985). Recall that in the standard benchmark case where consumption growth is i.i.d. and consumption equals dividends, the equity premium is given by $\gamma\sigma_c^2$. Thus, if the volatility of consumption growth equals 1% (see panel A of Table 1), even $\gamma = 60$ implies an equity premium of just 0.6%.

Since the aim of this article is not to address the equity premium puzzle, we choose to use the volatility of consumption growth measured over the longer sample 1929–2001, which is approximately 2.8%. Even this higher volatility of consumption growth is not sufficient to match the historical equity premium in the benchmark case with i.i.d. consumption growth and where dividends equal consumption, as in this case the implied equity premium is just 4.7%, still short of the historical average of 7.8% (see panel A of Table 1). In contrast, as panel B of Table 4 shows, the predictability induced by the labor income to consumption ratio raises the equity premium from 4.7 to 7.53%; a notable increase. Similarly, in the absence of any predictability, the volatility of returns would automatically be given by the volatility of consumption growth, 2.8%. Instead the variation in s_t^w almost triples the volatility of the returns on the market portfolio to 6.21%. Clearly, this increase in the volatility of returns is not enough to match the historical volatility of about 16% and this produces a Sharpe ratio that is higher in the calibration than its empirical counterpart.

In addition, notice that the level of the risk-free rate is a reasonable 3% and with a very small volatility.¹² Panel B also shows the mean and standard deviation of the share of labor income to consumption. The average value of the share is identical to \bar{s}^w , whereas the volatility and the correlation of ds_t^w and dC_t (in absolute terms) are slightly higher than the ones observed in empirical data (see panels A and B of Table 1), but not significantly so.

Predictability in artificial data. Table 5 reports the results of predictive regressions analogous to regressions (29)–(31) in our 10,000 years of simulated data.¹³ First, the line denoted regression 1 summarizes the univariate regression (29) in simulated data. The coefficient on s_t^w is negative as it was in the empirical regression. The magnitude of the coefficient, as well as the R^2 , is small compared to the empirical counterpart, something to be expected in our simulations, as explained below. Regression 2 shows the predictive results for the log dividend yield. Compared to the labor share s_t^w , the simulations show that the log dividend yield performs better in explaining future returns. Indeed,

Table 5
Forecasting regressions—simulations

		Forecasting Horizon			
		4	8	12	16
Regression 1	s_t^w	-0.055	-0.111	-0.170	-0.233
	Adjusted R^2	0.002	0.004	0.007	0.01
Regression 2	$\log(D/P)$	0.018	0.035	0.052	0.070
	Adjusted R^2	0.004	0.007	0.010	0.013
Regression 3	s_t^w	0.037	0.059	0.056	0.030
	$\log(D/P)$	0.027	0.049	0.066	0.077
	Adjusted R^2	0.004	0.007	0.010	0.013
Regression 4	s_t^w	-0.055	-0.111	-0.170	-0.233
	$(s_t^w)^2$	-0.251	-0.478	-0.670	-0.753
	Adjusted R^2	0.004	0.008	0.011	0.015

The table summarizes the result of the predictive regression

$$r_{t,t+K} = \alpha + \beta(K)\mathbf{x}_t + \varepsilon_{t+K}$$

in simulated data (see the notes to Table 4 for details of the simulation). Regressions 1–3 are as in Table 2. Regression 4 includes both the share of labor income to consumption, s_t^w , and the component of the share squared, $(s_t^w)^2$, that is orthogonal to the share itself.

¹² As is known, under power utility the risk-free rate is $r_t^f = \phi + \gamma E_t[dC_t/C_t] - 0.5\gamma(1 + \gamma)\sigma_c\sigma_c'$. The almost zero volatility of risk-free rate notwithstanding a $\gamma = 60$ just reflects the fact that in the model, $E_t[dC/C]$ is essentially constant, as discussed earlier.

¹³ We take the point estimates as population parameters due to the large size, 40,000 quarters, of our simulated sample.

regression 3 shows that when both the labor share s_t^w and log dividend yield are used as forecasting variables, the log dividend yield dominates.

To understand the simulation results, it is useful to turn to panel B of Figure 3, where the theoretical expected excess return on the market portfolio is plotted against s_t^w . As can be seen, nonlinearities are important and thus a linear specification—such as the one in regression 1—may miss predictable components in returns. Indeed, in regression 4, we also include $(s_t^w)^2$, orthogonalized to avoid multicollinearity problems. Now the R^2 doubles relative to the specification (29). The nonlinearity of expected return in share s_t^w also explains the better performance of the log dividend yield relative to s_t^w in simulations. From panel C of Figure 3, we see that the log dividend yield is nonlinear in the share of labor income to consumption, and thus it can better capture the nonlinearities in the expected excess return of the market portfolio.

In all specifications, the R^2 s are low compared to the data. The reason is that a model with power utility implies too little variation in expected excess returns as s_t^w fluctuates over time. Note that from panel B of Figure 3, the expected excess return increases at most by 2.5% (from 6 to 8.5%) for a drop in s_t^w from 0.95 to 0.7. This implies a regression coefficient of about -0.1 , which is similar to the numbers in Table 5. A reasonable conjecture is that, other type of preferences, such as habit formation, would result in a larger impact on expected excess returns following changes in the share of labor income to consumption. Our choice of the power utility is intended to focus exclusively on the channel linking s_t^w to expected excess returns without contaminating it with other effects, such as time-varying risk preferences.

3.3 Labor income and cross-sectional predictability

3.3.1 Cross-sectional implications. The starting point of our tests of the conditional CAPM is the beta representation given in expression (28). This conditional beta representation shows that both $\beta^{w,i}(\mathbf{s}_t)$ and $\beta^{M,i}(\mathbf{s}_t)$ are complex functions of all the state variables, namely, (s_t^1, \dots, s_t^n) . In what follows, though, we approximate the betas as linear functions of s_t^w rather than the whole distribution of shares. There are two reasons to focus on this approximation. First, we are interested in the role of labor income in asset pricing tests and it seems reasonable to initially concentrate on s_t^w . A second and more subtle reason is that conditioning by the distribution may not be necessary if the sorting procedure to generate the set of test portfolios correlates with the relative share \bar{s}^i/s_t^i . That is, the sorting itself takes into account the dependence of the loading on the relative share [see Equation (27)]. To see this more clearly, consider the expression for the price–dividend ratio, Equation (21). If the sorting procedure depends on the (normalized) price of the asset, such as size

and market-to-book, one is, according to the model, implicitly sorting by the relative share \bar{s}^i/s_t^i . The component, not captured by the sorting procedure is precisely s_t^w , and thus, it is enough to condition the CAPM on this variable alone. Our simulations below confirm that quantitatively this approximation is in fact what is needed to improve over the unconditional CAPM. Thus, we write the betas in Equation (28) as

$$\beta^{w,i}(\mathbf{s}_t) \approx \beta_1^{w,i} + \beta_2^{w,i} s_t^w \quad \text{and} \quad \beta^{M,i}(\mathbf{s}_t) \approx \beta_1^{M,i} + \beta_2^{M,i} s_t^w.$$

In this case Equation (28) becomes

$$E_t[dR_t^i] = \beta_1^{w,i} E_t[dR_t^w] + \beta_2^{w,i} E_t[s_t^w dR_t^w] + \beta_1^{M,i} E_t[dR_t^M] + \beta_2^{M,i} E_t[s_t^w dR_t^M]$$

We can condition down this expression to obtain

$$E[dR_t^i] = \beta_1^{w,i} E[dR_t^w] + \beta_2^{w,i} E[s_t^w dR_t^w] + \beta_1^{M,i} E[dR_t^M] + \beta_2^{M,i} E[s_t^w dR_t^M] \quad (32)$$

which is the version of the conditional CAPM that we test below.¹⁴ Tests of Equation (32) require an estimate of the return to human capital, which is not observable. We follow Jagannathan and Wang (1996) and Lettau and Ludvigson (2001a) and proxy the return on human capital as the growth rate of wages $\Delta \log(w_t)$.

3.3.2 Empirical results. Panel A of Table 6 reports tests on several empirical specifications that are consistent with the implications of the model. For every specification we report the estimate, the t -statistic, and the Shanken (1992) corrected t -statistic (in brackets). We use excess returns throughout, but, as is standard, we include an intercept in the different specifications even when it is not implied by the theory. The last column reports the R^2 and, below it, the adjusted R^2 (in brackets). We report the R^2 simply to provide an intuitive measure of the cross-sectional fit and to ease comparison with other studies but clearly the success of the model cannot be judged by “how high or low” is the R^2 .

The first line of panel A of Table 6 reproduces a familiar result, namely, the inability of the CAPM to explain the cross-section of returns of size and book-to-market sorted portfolios. The beta on the value-weighted return is not statistically significant, enters with the wrong sign,¹⁵ and the R^2 is a puny 8%. In addition, the intercept is strongly significant. Panel A

¹⁴ Numerical computations show that the betas are indeed nonlinear in s_t^w . Still, when we add nonlinear terms in the regressions below they turn out to have a negligible effect on the results. Moreover, this is true independently of whether we use empirical or artificial data. These nonlinearities then seem unimportant from either the theoretical or the empirical point of view.

¹⁵ The negative sign on the market portfolio is pervasive in the literature. See, for instance, Jagannathan and Wang (1996) and Lettau and Ludvigson (2001b).

Table 6
CAPM Fama–MacBeth regressions

	Constant	R^M	$s^w R^M$	R^w	$s^w R^w$	R^2 (Adjusted)
Panel A: s^w as a scaling variable						
1	3.45*	-0.88				8%
t -statistics	(3.75)	(-0.83)				[4%]
	[3.73]	[-0.82]				
2	3.41*	-0.84		-0.18		8%
t -statistics	(3.72)	(-0.79)		(-0.61)		[0.4%]
	[3.66]	[-0.78]		[-0.60]		
3	3.03*	-1.10	0.27*			54%
t -statistics	(3.30)	(-1.04)	(3.07)			[50%]
	[2.54]	[-0.85]	[2.40]			
4	2.65*	-0.21		-35	-0.06*	22%
t -statistics	(2.93)	(-0.19)		(-1.24)	(-2.74)	[10%]
	[1.88]	[-0.13]		[-0.82]	[-1.77]	
5	2.98*	-1.06	0.27*	-0.29	0.00	57%
t -statistics	(3.34)	(-1.03)	(3.13)	(-1.01)	(0.28)	[48%]
	[2.48]	[-0.83]	[2.37]	[-0.76]	[0.21]	
	Constant	R^M	s^w			R^2 (Adjusted)
Panel B: s^w as a factor						
6	5.57*	-3.22*	0.04*			45%
t -statistics	(4.49)	(-2.51)	(2.86)			[40%]
	[2.76]	[-1.64]	[1.77]			

Panel A: estimates of cross-sectional Fama–MacBeth (1972) regressions using the 25 Fama–French portfolios. In parentheses we report the uncorrected Fama–MacBeth (1972) t -statistic. The Shanken (1992) corrected t -statistics are reported in brackets. The unadjusted and adjusted (in brackets) R^2 are reported in the last column. $s^w R^M$ denotes the component of $s^w R^M$ orthogonal to R^M . Similarly, $s^w R^w$ denotes the component of $s^w R^w$ orthogonal to R^w . Panel A reports the specifications that are supported by the model introduced in Section 1. s^w is defined as labor income divided by consumption. Panel B: results where the variable s^w is introduced independently. Data are quarterly and the sample period is 1948–2001. *significance at the 5% level using the Fama–MacBeth standard errors.

of Figure 4 gives a visual impression of the poor performance of the CAPM: fitted and realized average returns fail to align along the 45° line. Line 2 of Table 6 summarizes that adding the excess return on labor income, R^w does not help much. A better definition of the market portfolio does not seem to improve the dismal performance of the CAPM.¹⁶

Next, we include in line 3, the interaction term $s^w R^M$, that is, $\beta_2^{M,i}$ in Equation (32) is different from zero. Because s^w is slow moving, the joint presence of $s^w R^M$ and R^M can result in severe multicollinearity problems. For this reason, we include only the component of $s^w R^M$ that is orthogonal to R^M . As shown in line 3, conditioning market returns by the variable s^w dramatically improves the cross-sectional fit to an adjusted R^2 of 50%. The coefficient on $s^w R^M$ is strongly significant and positive. To

¹⁶ This result stands in contrast to those of Jagannathan and Wang (1996) and Lettau and Ludvigson (2001b), who find that return to labor income is significant and it greatly increases the fit as measured by the R^2 . These two papers use a different sample and a dating convention advocated by Jagannathan and Wang, which lags labor income by one month. We run the regression in line 2 of Table 6 by lagging labor income by one quarter, and indeed find a significant positive coefficient in the sample 1963–1998, but still not in the 1948–2001 sample.

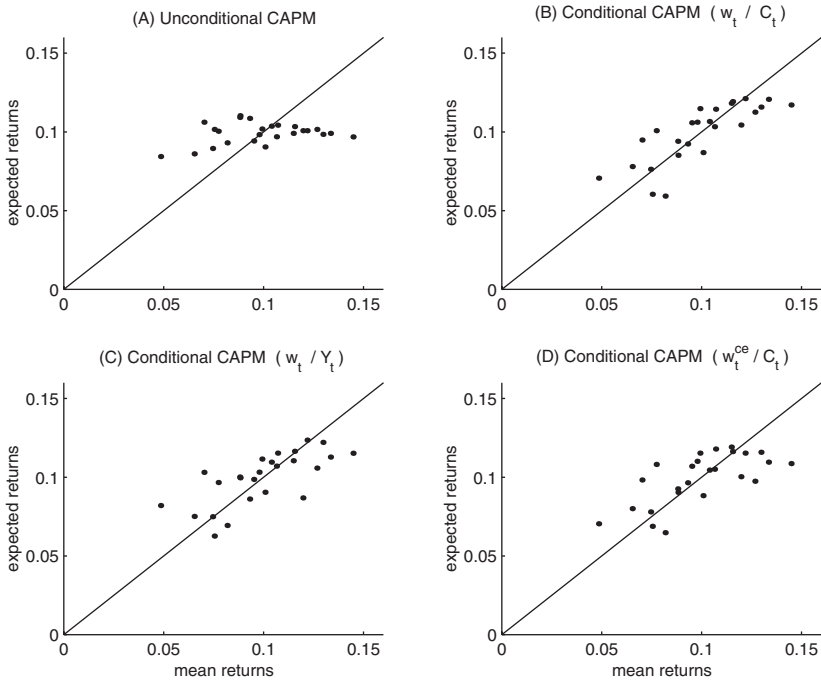


Figure 4
The conditional CAPM

Realized versus fitted returns on the 25 Fama and French portfolios. Panel A: unconditional CAPM. Panel B: conditional CAPM where the term $s_t^w R_t^M$ is included, and s_t^w is equal to labor income w_t divided by consumption C_t . Panel C: as in panel B but now s_t^w is equal to labor income w_t divided by disposable income Y_t . Panel D: as in panel B but now s_t^w is equal to compensation of employees w_t^{ce} divided by consumption C_t . Data are quarterly and the sample is 1948–2001. Average returns on the axis are annualized.

grasp visually the improvement over the unconditional CAPM, panel B of Figure 3 shows the corresponding plot of the fitted versus realized returns which now line up better along the 45° line than in the unconditional CAPM case. Although we do not report the results here in the interest of space, the improvement in the cross-sectional fit is robust to the alternative constructions of s_t^w introduced above. This can be seen in panels C and D where we plot fitted versus average returns when s_t^w is defined as w_t^{ce}/C_t and w_t/Y_t , respectively. Independent of the definition of s_t^w used then, the inclusion of the interaction term $s^w R^M$ helps in pricing these portfolios.

Table 6 also includes t -statistics using the correction in the computations of the standard errors proposed by Shanken (1992), which takes into account the fact that the betas are themselves estimated.¹⁷ In particular,

¹⁷ For a review of this point, see Cochrane (2001, page 239).

Shanken (1992) showed that under the assumption that asset returns have a conditional joint distribution with constant covariance matrix, the Fama–MacBeth procedure overstates the precision of the estimated parameters. As Lettau and Ludvigson (2001b) note, the size of the correction is much larger when using macroeconomic factors than when using purely financial variables. Indeed, notice that for the standard CAPM, there is essentially no difference between the uncorrected and the corrected t -statistic associated with R^M , whereas there is a much stronger correction when the share of labor income to consumption enters into the different specifications.¹⁸ It is reassuring then that the statistical significance of the coefficient on $s_t^w R_t^M$ is not affected by the Shanken correction.

In line 4, we test whether labor income per se is enough to improve the fit. Specifically, we drop the interaction term $s^w R^M$ and instead include the excess return on human wealth, R^w , plus the interaction term $s^w R^w$. Only the coefficient on the latter term is significant, but it enters with the wrong sign. Notice though that the significance of the coefficient on $s^w R^w$ is no longer there when the Shanken corrected t -statistics are used. Also, the improvement in the fit over the unconditional CAPM is much lower than in specification 3. Moreover, when the interaction term $s^w R^M$ is reintroduced in the full specification, line 5, $s^w R^w$ loses its significance, whereas that of $s^w R^M$ remains and the coefficient is identical to the one in line 3. We show in the simulations below that this is exactly in line with the predictions of the model.

Table 7 details the average pricing errors for the set of 25 portfolios associated with the different specifications tested above, where $S_i B_j$ denotes the portfolio whose size is in the i th quintile and book-to-market is in the j th quintile. For instance, as can be seen in the very first column, which details the average pricing errors associated with the standard CAPM, the small growth portfolio (S1B1) has a negative average pricing error (−0.8971), whereas the small value portfolio (S1B5) has a positive one (1.1993). The same pattern can be observed in the rest of the sizes, except the largest one. This is the value-spread puzzle. At the bottom of the table we report the result of a chi-square test of the null hypothesis that the pricing errors are *jointly* equal to zero and the corresponding p -value. It is worth recalling that the test is an asymptotic one and that its small sample properties are poor.¹⁹ Still, one can see that the null hypothesis that the pricing errors are jointly zero is rejected for the unconditional

¹⁸ However, the Shanken (1992) corrected t -statistics should be interpreted with caution. As Jagannathan and Wang (1998) show, “the standard errors obtained from the Fama–MacBeth procedure need not necessarily overstate the precision of estimates,” whenever the assumption of conditional homoskedasticity is violated.

¹⁹ For a clear introduction of the chi-square test, see Cochrane (2001, chap. 12). For the small sample properties of this test, see Hansen, Heaton, and Yaron (1995).

Table 7
Pricing errors

Port.	Regression line (see Table 6)					
	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6
S1B1	-0.8971	-0.8003	-0.5748	-0.5205	-0.3853	-1.1640
S1B2	0.2569	0.2713	0.2763	0.4048	0.2971	0.3128
S1B3	0.2385	0.2909	-0.2950	-0.1153	-0.1866	0.3695
S1B4	0.7808	0.7801	0.3089	0.7976	0.2857	0.6062
S1B5	1.1993	1.2366	0.7039	1.2322	0.7573	0.7463
S2B1	-0.5174	-0.4970	-0.3554	-0.4832	-0.3083	-0.4662
S2B2	0.0213	-0.0196	-0.2096	-0.4355	-0.2741	0.0079
S2B3	0.3949	0.3544	0.0315	0.2672	-0.0599	0.2056
S2B4	0.5264	0.4864	0.2525	0.5453	0.1565	0.1671
S2B5	0.8602	0.8809	0.5518	0.7950	0.5855	0.2052
S3B1	-0.3755	-0.4254	0.0839	-0.3583	-0.0005	-0.1244
S3B2	-0.0124	-0.0393	-0.1112	-0.2533	-0.1551	-0.1008
S3B3	-0.0676	-0.0751	-0.3722	-0.0965	-0.3990	-0.2208
S3B4	0.3020	0.2874	0.1176	0.5637	0.0620	-0.0987
S3B5	0.6278	0.5889	0.5252	0.2613	0.4644	0.4767
S4B1	-0.2823	-0.3589	0.5560	-0.1962	0.4283	0.5262
S4B2	-0.5771	-0.6023	-0.9665	-0.6752	-1.0291	-0.0003
S4B3	0.0073	-0.0248	0.0510	-0.1490	-0.0033	-0.0180
S4B4	0.0668	0.0499	-0.0807	0.2236	-0.1324	-0.1146
S4B5	0.4730	0.5208	0.2640	0.4800	0.3533	0.3411
S5B1	-0.6561	-0.6628	0.1250	-0.3685	0.1262	0.2003
S5B2	-0.8980	-0.9161	-0.5238	-0.5883	-0.5654	-0.7973
S5B3	-0.5537	-0.5281	0.0526	-0.2649	0.1130	-0.3325
S5B4	-0.5282	-0.5268	-0.0647	-0.6031	-0.0366	-0.4371
S5B5	-0.3898	-0.2713	-0.3461	-0.4629	-0.0940	-0.3002
Mean squared error	0.306529	0.304561	0.152421	0.261657	0.144556	0.18457
chi-square	65.3262	48.7303	39.0028	20.8030	27.8152	24.7798
p-value	0.0000	0.0013	0.0198	0.5329	0.1455	0.3617

Average pricing errors (in percentage) from the Fama–MacBeth (1972) regressions reported in Table 6, panels A and B, for each of the 25 Fama–French portfolios. S1 denotes the portfolio with the smallest firms and S5 the largest. Similarly, B1 includes the firms with the lowest book-to-market portfolios and B5, firms with the highest. The last three lines report the square root of average squared pricing error across all portfolios, the chi-square statistic for a test that the pricing errors are zero, and, finally, the corresponding *p*-value. Data are quarterly and the sample period is 1948–2001.

CAPM. This result holds as well for the case where the return on the human capital asset is included (line 2).

Lines 3–6 show the average pricing errors associated with the conditional CAPM. For instance, recall that in line 3 of Table 6 we showed that the beta on $s^w R^M$ is a statistically significant determinant of the cross-section of stock returns. Table 7 shows the improvements in the average pricing errors of, for example, S1B1 and S1B5. The value-spread is less of a puzzle for the conditional model. The chi-square drops significantly when compared to the unconditional version of the CAPM and we fail to reject the null hypothesis that the pricing errors are jointly equal to zero at the 1% level, although we can still reject at the 5%. The full specification, however, cannot be rejected at conventional statistical levels.

We have seen then that, scaling R^M by the share of labor income to consumption helps align the 25 Fama–French portfolios. Does s^w by itself, that is, as a factor, align these portfolios? Line 6, in panel B of Table 6, shows the Fama–MacBeth regression when s^w enters as a single factor. Although our tests, as shown in Equation (32), do not imply this regression, the results are nonetheless interesting to understand the effect of the labor-to-consumption ratio on the cross-section of stock returns. The loading in this factor is a statistically significant determinant of the cross-section and the adjusted R^2 jumps to 40% against 8% for the unconditional CAPM. Now the market portfolio is significant but enters with the wrong sign.

Finally, note that although the labor income share s_t^w enters significantly as a conditioning variable and helps explain a large cross-sectional variation in returns in the 25 Fama–French portfolios, the intercept in the cross-sectional regression test is still significantly different from zero. This finding is not uncommon in models where macrovariables are used as conditioning variables and a similar result can be found, for instance, in Jagannathan and Wang (1996) and Lettau and Ludvigson (2001b).²⁰ In Section 3.3.4 we turn to simulations to show to what extent, in the context of our model, misspecification can lead to large pricing errors and at the same time to large R^2 s. For example, when only $s_t^w R_t^M$ is used in the Fama–MacBeth procedure, our simulations show that the intercept is positive though the R^2 is 99%.²¹

3.3.3 Comparison with alternative models. How does our model compare with other asset pricing models? Table 8 summarizes similar regressions to the ones run in the previous section, using the consumption to wealth ratio of Lettau and Ludvigson (2001a, 2001b), cay , and the log dividend yield of the market portfolio, $\ln(D/P)$ as conditioning variables. The last set of regressions shows the performance of the Fama and French (1993) three-factor model. For each of these models, we report a basic regression, focusing on the performance of the alternative model itself and a second one where $s_t^w R_t^M$ is included. Our purpose with this last regression is to assess to what extent our variable survives the inclusion of variables that have been shown to perform well in the cross-section of stock returns.

²⁰ As Lettau and Ludvigson (2001b, page 1259) emphasize “Although the (C)CAPM can explain a substantial fraction of the cross-sectional variation in these 25 portfolios returns, this result suggests that the scaled models do a poor job of simultaneously pricing the hypothetical zero-beta portfolio.”

²¹ Since the initiation of the editorial process, new data have become available. We have run some additional checks with a sample extending to 2003. Results regarding the cross-section are essentially unaltered. As for the time-series predictability, the results remain also essentially unchanged when s_t^w is defined as either w_t^{cc}/C_t or w_t^l/Y_t , but weaken when defined as labor income divided by consumption.

Table 8
Comparison with alternative models

	Constant	R^M	$s^w R^M$	$cayR^M$	$\ln(D/P)R^M$	SMB	HML	R^2 [Adj]
1	0.33	1.63		0.07*				37%
<i>t</i> -statistics	(0.28)	(1.21)		(2.87)				[31%]
	[0.22]	[1.00]		[2.34]				
2	2.11*	-0.38	0.24*	0.02				60%
<i>t</i> -statistics	(2.40)	(-0.37)	(2.88)	(1.02)				[55%]
	[1.93]	[-0.37]	[2.35]	[0.84]				
3	0.61	1.31			2.34*			57%
<i>t</i> -statistics	(0.50)	(0.99)			(2.93)			[53%]
	[0.41]	[0.84]			[2.47]			
4	1.47	0.39	0.19*		2.21*			60%
<i>t</i> -statistics	(1.30)	(0.32)	(2.77)		(2.82)			[55%]
	[1.06]	[0.27]	[2.31]		[2.35]			
5	2.85*	-0.91				0.34	1.17*	72%
<i>t</i> -statistics	(2.48)	(-0.72)				(0.87)	(3.00)	[68%]
	[2.41]	[-0.70]				[0.87]	[3.00]	
6	2.74*	-0.81	0.04			0.34	1.17*	72%
<i>t</i> -statistics	(2.30)	(-0.62)	(0.51)			(0.89)	(3.00)	[67%]
	[2.24]	[-0.61]	[0.50]			[0.89]	[3.00]	

Estimates of cross-sectional Fama–MacBeth (1972) regressions using the 25 Fama–French portfolios for alternative pricing models. In parentheses, we report the uncorrected Fama–MacBeth *t*-statistic. The Shanken (1992) corrected *t*-statistics are reported in brackets. $s^w R^M$ denotes the component of $s^w R^M$ orthogonal to R^M . The unadjusted and adjusted (in brackets) R^2 are reported in the last column. Lines 1 and 2: comparison with the model of Lettau and Ludvigson (2001b). *cay* denotes the consumption to wealth ratio and $cayR^M$ denotes the interaction of this variable with the returns on the market portfolio. Data on *cay* were downloaded from Martin Lettau’s website (<http://pages.stern.nyu.edu/mlettau/>) and the sample period is 1952–2001. Lines 3 and 4: comparison with an asset pricing model where the log of the dividend yield, $\ln(D/P)$, is used as a conditioning variable. Lines 5 and 6: comparison with the Fama and French (1993) three-factor model. SMB and HML are the returns on the Fama–French mimicking portfolios related to size and book-to-market. Data is quarterly and, for lines 3–6, the sample period is 1948–2001. *statistical significance at the 5% level using the Fama–MacBeth (1972) standard errors.

Lines 1 and 3 show that both the consumption to wealth ratio and the dividend yield work well as conditioning variables when used to explain the cross-section of returns in the 25 Fama–French portfolios. These variables, when interacted with the market portfolio, are both significant and have intercepts that are not statistically different than zero.²² Lines 2 and 4 though show that the variable $s_t^w R_t^M$ is not “driven out” in either model, whereas $cay_t R_t^M$ is. The coefficient remains stable and strongly significant, independent of whether the Shanken correction is used or not.

Finally, line 5 reports the familiar performance of the Fama–French factors in this set of test portfolios. HML, the high-minus-low portfolio constructed as in Fama and French (1993) is strongly significant and, moreover, as shown in line 6 it drives out $s_t^w R_t^M$. The variable survives the

²² In the case of *cay*, this result is different than the one originally reported in Lettau and Ludvigson (2001b). They obtained an estimate for the rate of the zero beta portfolio that was too high when compared to the riskless borrowing rate [see line 5 of Table 1, in Lettau and Ludvigson (2001b, page 1256)].

inclusion of competing “interacting” variables like *cay* and $\ln(D/P)$, but not when included together with the Fama–French factors.

3.3.4 Simulation results. In this section, we use our simulations to investigate the model’s ability to reproduce the results in the previous section, that is, the “flat” relation between average and fitted excess returns of the unconditional CAPM and the improvement that results when conditioning by the share of labor income to consumption. In addition, these simulations allow us to assess the quantitative importance of approximating betas as a linear function of s_t^w .

Recall that we generated 10,000 years of artificial data for 200 identical firms, making use of the share process parameters in Table 4. Following Fama and French (1992), at the beginning of every simulated year we form 20 portfolios by sorting stocks based on their price–dividend ratios.²³ Notice that, given Equation (21), sorting by price–dividend ratios is, in the context of our model, the same as sorting by expected dividend growth and thus by expected returns. Panel D of Figure 5 plots the average price–dividend ratio of the different artificial portfolios versus their average returns. As can be seen, the sorting procedure in simulated data can generate a sizable cross-sectional spread in average returns, even when there is no unconditional cross-sectional dispersion in expected excess returns across individual assets. These portfolios then are an ideal testing ground for the implications of the model and for drawing comparisons with their empirical counterparts.

Table 9 summarizes the results of the tests of the CAPM and its conditional versions implied by the model in artificial data. Line 1 summarizes the regression that corresponds to the unconditional CAPM. In our model, the correlation with the market explains a larger percentage of the cross-sectional variation in average returns than in the empirical data. This is to be expected as, after all the standard deviation in the share of labor income to consumption is relatively small and thus so are the corresponding fluctuations in beta. It is rather surprising then that the R^2 is *only* 75%. Moreover, the intercept in the unconditional CAPM regression is strongly positive,²⁴ which shows that, in the context of our model, the unconditional CAPM is not performing well. A visual impression of the performance of the unconditional CAPM can be obtained by turning to panel A of Figure 5, where we show the CAPM fitted versus average returns in simulated data. Strikingly the model can reproduce the “flat” relation between average and fitted returns that is standard in this

²³ There is no room in our model for “size” effects as all assets are otherwise identical. A similar sorting procedure in simulated data has been done recently by Lettau and Wachter (2004).

²⁴ Given the size of the simulated sample, which is 40,000 quarters, we take the estimated coefficients to be population values and thus do not report the t -statistics.

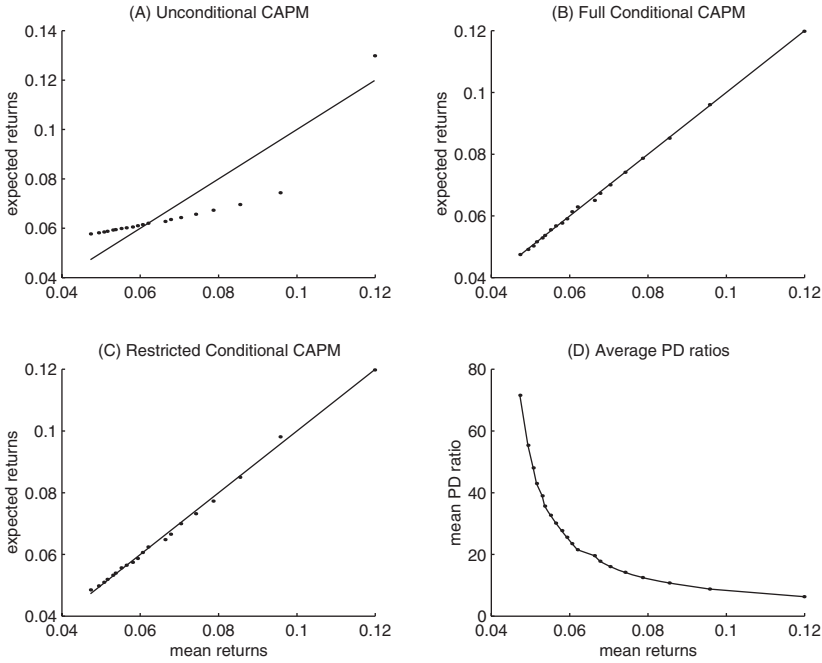


Figure 5

The conditional CAPM in simulations

Panels A–C: realized versus fitted returns on 20 portfolios sorted by price–dividend ratio in simulated data. Panel A: unconditional CAPM. Panel B: conditional CAPM for the full specification, where all the terms $s_t^w R_t^M$, R_t^w and $s_t^w R_t^w$ are added to the unconditional CAPM. Panel C: conditional CAPM for the restricted specification, where only the interaction term $s_t^w R_t^M$ is added to the unconditional CAPM. Panel D: average price–dividend ratios of the 20 portfolios plotted against their average returns. All entries in the axis are annualized.

set of test portfolios. The only exception to this is the portfolio with the highest expected excess return, which lies almost on the 45° line. This is unsurprising. This portfolio correlates strongly with the market, since the assets in this portfolios are those for which the current contribution to the total amount of dividends paid by the market portfolio is very large. Thus, they co-move with the market portfolio relatively more than the other assets, and thus, the unconditional CAPM can price this asset better than others.

Line 2 of Table 9 summarizes the case where the interaction term $s^w R^M$ is included. The introduction of this term increases the R^2 to almost 100% and thus clearly improves the cross-sectional fit. Comparing panel A in Figure 5 with panel C, we see that the introduction of the interaction term aligns the portfolios along the 45° line almost perfectly. In addition, the size of the coefficient on $s^w R^M$, 0.25, is remarkably close to its empirical counterpart. The intercept of this model is positive, though only 1/3 of the

Table 9
CAPM Fama–MacBeth regressions—simulations

	Constant	R^M	$s^w R^M$	R^w	$s^w R^w$	$R^2(\text{Adjusted})$
Panel A: s^w as a scaling variable						
1	1.29	0.71				75.81%
2	0.42	0.82	0.25			99.75%
3	0.00	1.78	0.10	0.81	0.17	99.93%
	Constant	R^M	s^w			$R^2(\text{Adjusted})$
Panel B: s^w as a factor						
4	1.16	0.62	13.53			75.96%
	Constant	R^M	$s^w R^M$	HML		$R^2(\text{Adjusted})$
Panel C: HML						
5	0.08	1.83		0.86		99.25%
6	0.33	1.07	0.20	0.32		99.81%

Panel A: estimates of cross-sectional Fama–MacBeth (1972) regressions in simulated data (see the notes to Table 4 for details of the simulation). From the simulated price–dividend ratios and returns of the 200 firms, we form 20 portfolios sorted by price–dividend ratio at the beginning of each year, and compute quarterly time series of value-weighted returns corresponding to each portfolio. The Fama–MacBeth (1972) cross-sectional regression procedure is applied to the 20 portfolios. The last column reports the unadjusted R^2 . $s^w R^M$ denotes the component of $s^w R^M$ orthogonal to R^M . Similarly, $s^w R^w$ denotes the component of $s^w R^w$ orthogonal to R^w . Panel A reports the specifications that are supported by the model introduced in Section 1. Panel B: results where the variable s^w is introduced independently. Panel C: results with a simulated HML mimicking portfolio. HML is constructed as the value-weighted return on a portfolio of “value” stocks, defined as those with a price–dividend ratio below the median, minus the value-weighted return of “growth” stocks, those with a price–dividend ratio above the median.

value of the corresponding one in the unconditional CAPM. This result shows that it is indeed possible to obtain a high R^2 in the cross-sectional fit, but retain a strongly positive intercept because of model misspecification. Indeed, for the full specification case (line 3) the intercept is zero. In terms of the R^2 , the improvement over the specification where only $s^w R^M$ is included is puny (compare panels B and C of Figure 5.) Notice also that the approximation made in Section 3.3.1 of beta as a linear function of s_t^w seems rather innocuous and that, in terms of cross-sectional fit, is not relevant.²⁵

Finally, in lines 5 and 6, we report the cross-sectional regression, where a simulated HML factor has been included. Obviously, we do not have a variable that is the counterpart to “book” in our model, and thus to construct the HML factor we proceed as follows. From the returns of the 200 firms, we compute the value-weighted returns of two portfolios that are constructed using stocks with price–dividend ratios below and above the median, R_t^{Value} and R_t^{Growth} , respectively. Then, we follow Fama

²⁵ The result that the full specification leads to a zero; Jensen’s alpha in simulation stands in contrast with the corresponding one in the data. As already mentioned, the finding of positive alphas is not uncommon in the literature, and, as seen in our simulations, it can be due to model misspecification.

and French (1993) and construct the returns on the HML portfolio as $R_t^{\text{Value}} - R_t^{\text{Growth}}$. As line 5 summarizes HML explains most of the cross-sectional variation of returns in simulated data. Moreover, the size of the coefficient is on the same order of magnitude of the point estimate of its empirical counterpart, 1.17, though it drops when $s^w R^M$ is included.²⁶

In summary, accounting for variation in betas due to fluctuations in the share of labor income to consumption, a purely macroeconomic variable, improves the cross-sectional fit of the 25 Fama–French portfolios over the unconditional version of the CAPM, and the model can produce magnitudes that are comparable to their empirical counterparts.

4. Conclusions

We have proposed a simple general equilibrium model to illustrate the effect that the inclusion of labor income has on asset pricing tests. We show that equilibrium expected returns change as the fraction of total income funded by labor income fluctuates over time. The reason is that, variations in this fraction affect the conditional covariance between equilibrium returns and consumption growth and, as a consequence, it results in variation in the premiums investors require to hold stocks. We then obtain a new and simple testable implication, namely, that the ratio of labor income to consumption should forecast stock returns. This is strongly confirmed in the data. The regression of stock returns on lagged values of this ratio produces statistically significant coefficients and adjusted R^2 s that are larger than those generated when using the dividend–price ratio as a single explanatory variable.

Our model also has implications for the cross-section of stock returns. In particular, we show that the asset's beta depends on both a proxy for expected dividend growth and the share of consumption funded by labor income. When we test these implications in the set of 25 portfolios sorted by size and book-to-market, we find that the version of the conditional CAPM advanced in this article performs better than the traditional unconditional CAPM. Moreover, we have shown, via simulations, that the model can reproduce the standard flat relation between average returns and fitted returns produced by the unconditional CAPM. Thus, ignoring the variation of the labor income to consumption ratio, even when this is small, can lead to severe misspricing in the cross-section. Deviations of labor income from its long run relation with consumption then contain useful information about returns in both the time series and the cross-section.²⁷

²⁶ In our framework, other conditioning variables such as *cay* and $\ln(D/P)$ are nonlinear transformations of the share of labor income to consumption. We omit then reporting cross-sectional results, using these variables as they are essentially identical to lines 1–4 in Table 9.

²⁷ Recently, Lustig and Verdelhan (2004) have used the version of the conditional consumption CAPM advanced in this article to address cross-sectional differences in average returns in a set of currency-sorted portfolios.

There are several natural candidates as extensions of the present analysis. One possible extension is to allow for richer preference specifications than the ones contemplated here. Our share technology can be coupled with a more sophisticated stochastic discount factor, such as a habit persistence one, to add quantitatively to the effects discussed in this article and obtain premiums and volatility of returns closer to those observed in the data. Still, the present article shows that relatively small variations in the share of consumption funded by labor income can have a relatively large impact on asset prices and returns. Another, more demanding, extension would be to allow for the possibility of an endogenous labor supply.²⁸ Of course, such models would also have to explain patterns in the labor supply, not only on consumption and returns, and in this they represent a much more challenging direction.

Appendix

Proof of Proposition 1. From the pricing formula, we obtain immediately

$$P_t^i = C_t^\gamma E_t \left[\int_t^\infty e^{-\phi(\tau-t)} C_\tau^{1-\gamma} s_\tau^i d\tau \right].$$

For all $i = 1, \dots, n$, define

$$Y_t^i = e^{-\phi t} C_t^{1-\gamma} s_t^i.$$

An application of Itô's lemma shows

$$dY_t^i = \left(\left(-\phi + (1-\gamma)\mu_{c,t} - \frac{1}{2}\gamma(1-\gamma)\sigma_c' \sigma_c + (1-\gamma)\theta^i \right) Y_t^i + \sum_{j=1}^n Y_t^j \lambda_{ji} dt - (1-\gamma) Y_t^i \sum_{j=1}^n s_t^j \theta^j \right) dt + ((1-\gamma) Y_t^i \sigma_c' + Y_t^i \sigma^i(s)) d\mathbf{B}_t.$$

Recalling the assumption $\mu_{c,t} = \bar{\mu}_c + \sum_{i=1}^n s_t^i \theta^i$, we obtain the convenient formula

$$dY_t^i = \left(-\tilde{\phi}^i Y_t^i + \sum_{j=1}^n Y_t^j \lambda_{ji} dt \right) dt + ((1-\gamma) Y_t^i \sigma_c' + Y_t^i \sigma^i(s)) d\mathbf{B}_t,$$

where $\tilde{\phi}^i = \phi - (1-\gamma)\bar{\mu}_c + \frac{1}{2}\gamma(1-\gamma)\sigma_c' \sigma_c - (1-\gamma)\theta^i$. We can rewrite the process in vector form

$$d\mathbf{Y}_t = -\mathbf{A} \cdot \mathbf{Y}_t dt + \Sigma(\mathbf{Y}) \cdot d\mathbf{B}_t,$$

where $\mathbf{A} = \mathbf{I}(\phi) - \mathbf{\Lambda}$. Using the expectation operator, we then find

$$E_t[\mathbf{Y}_\tau] = \Phi(\tau - t) \cdot \mathbf{Y}_t,$$

²⁸ Boldrin, Christiano, and Fisher (2001) propose a two-sector RBC model where agents have habit persistence preferences over consumption and endogenous labor supply to account for standard moments of interest in this literature, including the equity premium, the Sharpe ratio, and the level of the risk-free rate. See, also, Danthine and Donaldson (2002), Lettau (2003), and Wei (2003).

where $\Phi(\tau - t) = \mathbf{U}e^{-\omega(\tau-t)}\mathbf{U}^{-1}$, and $[e^{-\omega(\tau-t)}]$ is the diagonal matrix with $e^{-\omega_i(\tau-t)}$ on its i th element, where ω_i is the i th eigenvalue of \mathbf{A} , and where \mathbf{U} is the matrix with \mathbf{A} eigenvectors on its columns. Thus, for asset i , we have

$$E_t \left[e^{-\phi(\tau-t)} C_\tau^{1-\gamma} s_\tau^j \right] = E_t [Y_\tau^i] = \sum_{j=1}^n \Phi_{i,j}(\tau) Y_\tau^j = C_t^{1-\gamma} \sum_{j=1}^n s_t^j \left(\sum_{k=1}^n [\mathbf{U}]_{i,k} [\mathbf{U}^{-1}]_{k,j} e^{-\omega_k(\tau-t)} \right)$$

Substituting in the pricing function

$$\begin{aligned} P_t^i &= C_t^\gamma E_t \left[\int_t^\infty e^{-\phi(\tau-t)} C_\tau^{1-\gamma} s_\tau^j d\tau \right] = C_t \sum_{j=1}^n s_t^j \left(\sum_{k=1}^n [\mathbf{U}]_{i,k} [\mathbf{U}^{-1}]_{k,j} \int_t^\infty e^{-\omega_k(\tau-t)} d\tau \right) \\ &= C_t \sum_{j=1}^n s_t^j \left(\sum_{k=1}^n \frac{[\mathbf{U}]_{i,k} [\mathbf{U}^{-1}]_{k,j}}{\omega_k} \right) = C_t \sum_{j=1}^n (\mathbf{e}_i \mathbf{U} \mathbf{\Omega}^{-1} \mathbf{U}^{-1} \mathbf{e}_j) s_t^j \\ &= \mathbf{b}'_i \cdot \mathbf{D}_t \end{aligned}$$

where $\mathbf{\Omega}$ is a diagonal matrix with \mathbf{A} 's eigenvalues on its principal diagonal, and

$$\mathbf{b}'_i = \mathbf{e}_i \mathbf{U} \mathbf{\Omega}^{-1} \mathbf{U}^{-1} = \mathbf{e}_i \mathbf{A}^{-1}, \quad (\text{A1})$$

where the last equality stems from the standard decomposition $\mathbf{A} = \mathbf{U} \mathbf{\Omega} \mathbf{U}^{-1}$. ■

Proof of Proposition 2. Given the pricing function, *Itô's* lemma immediately yields the diffusion component of the expected return process, defined by

$$dR_t^i = \frac{dP_t^i + D_t^i dt}{P_t^i} - r_t dt$$

This is given by

$$\sigma'_{R,i} = \sigma'_c + \frac{\mathbf{b}'_i \cdot \mathbf{I}(\mathbf{s}_t) \cdot (\mathbf{v}' - 1_n \mathbf{s}_t \cdot \mathbf{v}')}{\mathbf{b}'_i \cdot \mathbf{s}_t}$$

Under power utility, the stochastic discount factor is $m_t = e^{-\phi t} C_t^{-\gamma}$ which follows the process

$$\frac{dm_t}{m_t} = -r_t dt - \gamma \sigma'_c d\mathbf{B}_t$$

where $r_t = \phi + \gamma \mu_{c,t} - \frac{1}{2} \gamma (1 + \gamma) \sigma'_c \sigma_c$. Thus, the expected return of asset i is

$$E[dR_t^i] = \gamma \sigma'_{R,i} \sigma_c,$$

yielding Equation (16) as we recall that $\theta^i = \mathbf{v}'_i \cdot \sigma_c$. ■

Proof of Proposition 3. Under log utility, the pricing equation is immediate from

$$P_t^i = E_t \left[\int_t^\infty \frac{U_c(C_\tau, \tau)}{U_c(C_t, t)} D_\tau^i d\tau \right] = E_t \left[\int_t^\infty e^{-\phi(\tau-t)} \frac{C_t}{C_\tau} s_\tau^j C_\tau d\tau \right] \quad (\text{A2})$$

$$= C_t \int_t^\infty e^{-\phi(\tau-t)} E_t [s_\tau^j] d\tau = C_t \left\{ \frac{a s_t^j}{\phi(a + \phi)} + \frac{s_t^j}{a + \phi} \right\}, \quad (\text{A3})$$

where we used the fact that Equation (17) implies that $E_t[s_t^i] = \bar{s}^i + (s_t^i - \bar{s}^i)e^{-a(\tau-t)}$. The expression for the price–dividend ratio of asset i follows immediately from (A3). ■

Proof of Proposition 4. The result follows from a straightforward application of *Itô*'s lemma to the pricing function Equation (A3), together with the fact that under log utility $E_t[dR_t^i] = \text{cov}_t(dR_t^i, \frac{dC_t}{C_t})$ and $E_t[dR_t^M] = \text{cov}_t(dR_t^M, \frac{dC_t}{C_t})$. ■

Proof of Proposition 5. In the log economy, $E_t[dR_t^i] = \text{cov}_t(dR_t^i, dR_t^{TW})$, and thus $E_t[dR_t^{TW}] = \text{var}_t(dR_t^{TW})$. The beta representation then follows trivially from Equation (24). ■

Proof of Proposition 6. To obtain the version of the CAPM with respect to the market portfolio, define

$$P_t^{TW} = P_t^M + P_t^w \quad \text{and} \quad \Phi(\mathbf{s}_t) = \frac{P_t^w}{P_t^w + P_t^M},$$

where P_t^w is the price of the human capital asset, P_t^M is the price of the market portfolio, and, recall, $\mathbf{s}_t = (s_t^1, s_t^2, \dots, s_t^J)$. Then the return on the total wealth portfolio is given by

$$dR_t^{TW} = \Phi(\mathbf{s}_t)dR_t^w + (1 - \Phi(\mathbf{s}_t))dR_t^M, \quad (\text{A4})$$

where dR_t^w is the rate of return on the human capital asset. Given that

$$E_t[dR_t^i] = \text{cov}_t(dR_t^i, dR_t^{TW})$$

we can use the definition of the return on the total wealth portfolio, Equation (A6), to obtain

$$E_t[dR_t^i] = \Phi(\mathbf{s}_t)\text{cov}_t(dR_t^i, dR_t^w) + (1 - \Phi(\mathbf{s}_t))\text{cov}_t(dR_t^i, dR_t^M). \quad (\text{A5})$$

Then, Equation (37) implies that the conditional expected rates of return on both the human capital asset and the market are given by

$$[E_t[dR_t^w] E_t[dR_t^M]]' = \Sigma^{wM} [\Phi(\mathbf{s}_t)(1 - \Phi(\mathbf{s}_t))]', \quad (\text{A6})$$

where Σ^{wM} is the variance–covariance matrix of dR_t^w and dR_t^M . From Equation (38) we can get an expression for $[\Phi(\mathbf{s}_t)(1 - \Phi(\mathbf{s}_t))]'$ that we can readily substitute back into Equation (37) to obtain the beta representation Equation (28). ■

References

- Ang, A., and G. Bekaert, 2001, “Stock Return Predictability: Is it There?” Working paper, Columbia University.
- Ball, R., and S. P. Kothari, 1989, “Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns,” *Journal of Financial Economics*, 25, 51–74.
- Barberis, N., M. Huang, and T. Santos, 2001, “Prospect Theory and Asset Prices,” *Quarterly Journal of Economics*, 116(1), 1–53.
- Bodurtha, J. N., and N. C. Mark, 1991, “Testing the CAPM with Time-varying Risks and Returns,” *Journal of Finance*, 46(4 September), 1485–1505.
- Boldrin, M., L. J. Christiano, and J. D. Fisher, 2001, “Habit Persistence, Asset Returns, and the Business Cycle,” *American Economic Review*, 91(1 March), 149–166.
- Bollerslev, T., R. F. Engle, and J. M. Wooldridge, 1988, “A Capital Asset Pricing Model with Time-varying Covariances,” *Journal of Political Economy*, 96(1), 116–131.

- Bossaerts, P., and R. C. Green, 1989, "A General Equilibrium Model of Changing Risk Premia: Theory and Tests," *Review of Financial Studies*, 2(4), 467–493.
- Braun, P. A., D. B. Nelson, and A. M. Sunier, 1995, "Testing the CAPM with Time-varying Risks and Returns," *Journal of Finance*, 50(5), 1575–1603.
- Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–400.
- Campbell, J. Y., 1996, "Understanding Risk and Return," *Journal of Political Economy*, 104(2), 298–345.
- Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–251.
- Campbell, J. Y., and P. Perron, 1991, "Pitfalls and opportunities: What macroeconomists should know about unit roots," NBER Technical Working Paper, no. 100.
- Campbell, J. Y., and R. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.
- Campbell, J. Y., and R. Shiller, 1988b, "Stock Prices, Earnings, and Expected Dividends," *Journal Of Finance*, 43, 661–676.
- Cho, Y., and R. F. Engle, 1999, "Time-varying Betas and Asymmetric Effects of News: Empirical Analysis of Blue Chip Stocks," NBER Working Paper, no. 7330, September.
- Cochrane, J. H., 1996, "A Cross Sectional Test of an Investment Based Asset Pricing Model," *Journal of Political Economy*, 104(3), 572–621.
- Cochrane, J. H., 2001, *Asset Pricing*, Princeton University Press, Princeton, NJ.
- Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara, 2004, "Two Trees," Working paper, University of Chicago.
- Danthine, J. P., and J. Donaldson, 2002, "Labor Relations and Asset Returns," *Review of Economic Studies*, 69(1), 41–64.
- Duffee, G. R., 2004, "Time-variation in the covariance between stock returns and consumption growth," Working paper, Haas School of Business, University of California, Berkeley.
- Duffie, D., 1996, *Dynamic Asset Pricing Theory* (2d ed.), Princeton University Press, Princeton, NJ.
- Evans, M. D., 1994, "Expected Returns, Time-varying Risk, and Risk Premia," *Journal of Finance*, 49(2), 655–679.
- Fama, E. F., and J. D. MacBeth, 1972, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.
- Fama, E. F., and K. R. French, 1988a, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–27.
- Fama, E. F., and K. R. French, 1988b, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96, 246–273.
- Fama, E. F., and K. R. French, 1992, "The cross-section of Expected Returns," *Journal of Finance*, 47, 427–465.
- Fama, E. F., and K. R. French, 1989, "Business Conditions and Expected Stock Returns on Stocks and Bonds," *Journal of Financial Economics*, 25, 23–49.
- Fama, E. F., and K. R. French, 1993, "Common Risk Factors in the Returns of Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.
- Fama, E. F., and K. R. French, 1997, "Industry Costs of Equity," *Journal of Financial Economics*, 43, 153–193.

- Ferson, W. E., 1989, "Changes in Expected Security Returns, Risk, and the Level of Interest Rates," *Journal of Finance*, 44, 1191–1218.
- Ferson, W. E., and C. R. Harvey, 1991, "The Variation of Economic Risk Premiums," *Journal of Political Economy*, 99(2), 385–415.
- Ferson, W. E., and C. R. Harvey, 1999, "Conditioning Variables and the cross-section of Stock Returns," *Journal of Finance*, 54, 1325–1361.
- Ferson, W. E., S. Kandel, and R. F. Stambaugh, 1987, "Tests of Asset Pricing with Time-varying Expected Risk Premiums and Market Betas," *Journal of Finance*, 42(2), 201–220.
- Franzoni, F., 2001, "Where is Beta Going? The Riskiness of Value and Small Stocks," Working paper, MIT.
- Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
- Hansen, L. P., J. Heaton, and A. Yaron, 1995, "Finite Sample Properties of Some Alternative GMM Estimators," *Journal of Business Economic Statistics*, 4, 262–280.
- Harvey, C. R., 1989, "Time-varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics*, 24, 289–317.
- Hodrick, R. J., 1992, "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 357–386.
- Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51(1), 3–53.
- Jagannathan, R., and Z. Wang, 1998, "An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression," *Journal of Finance*, 53(4), 1285–1309.
- Keim, D. B., and R. F. Stambaugh, 1986, "Predicting Stock Returns in the Stock and Bond Markets," *Journal of Financial Economics*, 17, 357–390.
- Lamont, O., 1998, "Earnings and Expected Returns," *Journal of Finance*, 53, 1563–1587.
- Lettau, M., 2003, "Inspecting the Mechanism: The Determination of Asset Prices in the RBC Model," *The Economic Journal*, 113, 550–575.
- Lettau, M., and S. Ludvigson, 2001a, "Consumption, Aggregate Wealth, and Expected Returns," *Journal of Finance*, 55(3), 815–849.
- Lettau, M., and S. Ludvigson, 2001b, "Resurrecting the CCAPM: A Cross-Sectional Test of Linear Factor Models When Risk Premia are Time-Varying," *Journal of Political Economy*, 109(6), 1238–1287.
- Lettau, M., and S. Ludvigson, 2004, "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94(1), 276–299.
- Lettau, M., and J. Wachter, 2004, "Why is long-horizon equity less risky? A duration based explanation of the value premium," Working paper, New York University.
- Lewellen, J., and S. Nagel, 2003, "The conditional CAPM does not explain Asset Pricing Anomalies," Working paper, MIT.
- Longstaff, F., and M. Piazzesi, 2004, "Corporate Earnings and the Equity Premium," *Journal of Financial Economics*, 74(3), 401–421.
- Lustig, H., and A. Verdelhan, 2004, "The cross-section of Foreign Currency Risk Premia and US Consumption Growth Risk," Working paper, University of Chicago.
- Mayers, D., 1972, "Non-marketable Assets and Capital Market Equilibrium under Uncertainty," in M. Jensen (ed.), *Studies in the Theory of Capital Markets*, Praeger, New York, 223–248.
- Mehra, R., and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

- Menzly, L., T. Santos, and P. Veronesi, 2004, "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.
- Merton, R., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–887.
- Ng, L., 1991, "Tests of the CAPM with Time-varying Covariances: A Multivariate GARCH Approach," *Journal of Finance*, XLVI(4 September), 1507–1521.
- Ng, S., and P. Perron, 1995, "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lags," *Journal of the American Statistical Association*, XC(429), March, 268–281.
- Roll, R., 1977, "A Critique of the Asset Pricing Theory's Tests: Part I," *Journal of Financial Economics*, 4, 129–176.
- Rubinstein, M., 1976, "The Valuation of Uncertain Income Streams and the Price of Options," *Bell Journal of Economics*, 7, 407–425.
- Santos, T., and P. Veronesi, 2001, "Labor Income and Predictable Stock Returns," NBER Working Paper, no. 8309.
- Shanken, J., 1990, "Intertemporal Asset pricing: An Empirical Investigation," *Journal of Econometrics*, 45, 99–120.
- Shanken, J., 1992, "On the Estimation of Beta Pricing Models," *Review of Financial Studies*, 5(1), 1–33.
- Timmermann, A., 1993, "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 108, 1135–1145.
- Veronesi, P., 2000, "How does Information Quality Affect Stock Returns?" *Journal of Finance*, 55, 807–837.
- Wei, M., 2003, "Human Capital, Business Cycle, and Asset Pricing," Working paper, Graduate School of Business, Columbia University.