Labor Income and Predictable Stock Returns *

Tano Santos
Graduate School of Business
University of Chicago

Pietro Veronesi
Graduate School of Business
University of Chicago

This Version: June 15, 2000

*We thank Fernando Alvarez, Nick Barberis, John Y. Campbell, John H. Cochrane, George Constantinides, Kent Daniel, Gene Fama, Lars P. Hansen, John Heaton, Martin Lettau, Sydney Ludvigson and Lubos Pastor for comments and suggestions. We thank Lior Menzly for excellent research assistance. All errors are our own. Tano Santos e-mail: jesus.santos@gsb.uchicago.edu. Pietro Veronesi e-mail: pietro.veronesi@gsb.uchicago.edu.

Abstract

We propose and test a novel economic mechanism that generates time varying expected returns. In our model, investors' income has two sources, wages and dividends, that grow stochastically over time. As a consequence, the fraction of total income produced by wages changes over time depending on economic conditions. We show that as this fraction fluctuates, the risk premium that investors require to hold stocks varies as well.

Tests of the main implications of the model are strongly confirmed in the data. A regression of stock returns on lagged values of the labor income to consumption ratio produces statistically significant coefficients and adjusted R^2's that are larger than those generated when using the dividend price ratio. We also derive and test the model's implications for the cross section. We find considerable improvements on the performance of both the conditional CAPM and CCAPM when compared to their unconditional counterparts.

1 Introduction

The substantial predictability of stock returns is by now widely documented. The main finding is that variables like the price-dividend ratio and term premia can predict variation on stock returns at long horizons.[1] This evidence has survived a decade long effort to tackle many of the econometric issues that are relevant for evaluating these effects "at reasonable if not overwhelming levels of statistical significance,"(Campbell (2000), page 9.) Furthermore, predictability is not unique to the standard US data set, but it can be found in many other countries as well (see Campbell (1999), Table 12, panel B.)

Extant economic explanations for the predictability of stock returns rely either on investors' learning about some unobservable fundamental process (Timmerman (1993, 1996) and Veronesi (2000)) or cyclical variations in investors' risk aversion (Campbell and Cochrane (1999) and Barberis, Huang and Santos (2000)).

In this paper, we propose and test a different economic mechanism that generates time varying expected returns. In our set up investors' income has two sources, wages and dividends, that grow stochastically over time. As a consequence, the fraction of total income produced by wages changes over time depending on economic conditions. It is then intuitive that as this fraction fluctuates, the risk premium that investors require to hold stocks varies as well: For example, when investors' consumption is mainly financed by wages, the relationship between stock returns and consumption growth weakens thereby generating a lower required risk premium.

We formalize this intuition by describing a simple, pure-exchange Lucas (1978) economy where investors are endowed with a stochastic labor income stream along with their income from stock market investments. Under mild assumptions on the stochastic processes for dividends and labor income, we obtain closed form solutions for the equilibrium stock market price and solve for stock returns. We then show that expected

[1] A lucid exposition of the main findings can be found in Campbell, Lo, and MacKinley (1997, Chapter 7.) See also Campbell (2000) and Cochrane (2000).

returns vary over time with the ratio of labor income to consumption. We also obtain closed form solutions for long-term expected returns and prove that they depend on the current labor-income to consumption ratio. In addition, we also demonstrate that expected stock returns have an equivalent representation as a function of the dividend yield.

We also develop and test the cross sectional implications of our model. In particular, we show that expected returns on individual securities have a stochastic beta representation, where "beta" is itself a function of the labor-income to consumption ratio. Intuitively, when the labor income to consumption ratio increases, the expected returns for the whole market decreases, as mentioned earlier. However, it may decrease differently across securities thereby affecting their "beta" with respect to the market returns.

We test the main predictions of the model by running predictive regressions using as explanatory variables both the ratio of labor income over total consumption and/or the dividend price ratio, which are plotted in Figure 1. Our first main finding is that for the overall sample 1946-1999 the ratio of labor income to total consumption performs much better than the (log) dividend price ratio to forecast future returns. For example, the R^2 of the one year predictive regression is 6.1% (against 3.9% for the log dividend-price ratio) and it reaches 34.6% for the four year ahead regression (against 20.7% for the dividend price ratio.) Since this result may be driven by the inclusion of the 1995-1999 period in the data sample (which witnessed a low dividend price ratio and high returns), we also run regressions over the more conventional 1952-1994 data period where it is known that the dividend price ratio is working well (see e.g. Campbell et al. (1997)). In this case the dividend price ratio is performing better than the wages-to-consumption ratio, but the latter is still a significant regressor at long horizons where it still produces a good forecasting power: The R^2 is 3% for the one year ahead regression but it reaches 22.6% for the 4 year ahead regression.

More interestingly, the inclusion of both regressors – either linearly or in a multiplicative fashion – dramatically improves upon the predictive power of the regression at every horizon. For example, still in the 1952-1994 sample the (adjusted) R^2 ranges between 25.1% at the one year horizon to above 60% at the four year horizon. This finding is indeed fully consistent with our model, that implies that both expected returns and the dividend yield are non linear functions of the labor income to consumption ratio. It follows that the dividend yield may proxy for the non linear relation between expected returns and this ratio and it is then not surprising that it improves the predictability.

We next test the cross-sectional implications of the model, that is, the stochastic beta representation for the expected returns of individual securities, where the beta of each security is itself a function of the labor income to consumption ratio. We test whether conditioning the CAPM and the CCAPM by the share of labor income to consumption improves the fit on the cross section relative to their unconditional counterparts.

The portfolios we use to test this hypothesis are the 25 Fama and French (1993) portfolios, which are portfolios sorted by size and book to market. These portfolio allow us to test whether our conditioning variable can explain the justly celebrated value premium uncovered by Fama and French. In our sample, whereas the unconditional CAPM only explains around 1% of the cross section, its conditional version can explain as much as 43% of the cross section. Similar results are obtained for the conditional Consumption CAPM where the R^2 jumps from 5% in the unconditional version to an adjusted R^2 as high as 30.7%. The results then show that conditioning by the fraction of labor income to consumption considerable improves the fit of the CAPM and CCAPM and lends further support to the economic model provided in this paper.

Clearly we are not the first to note the importance of incorporating labor income in empirical asset pricing models and references on human capital and asset returns go as far back as Mayers (1972) and Fama and Schwert (1977). More recently, authors like Jagannathan and Wang (1996) and Campbell (1996) have included human capital returns to improve on the definition of the market portfolio. For instance, in tests of the cross section, Jagannathan and Wang (1996) find that the inclusion of human capital in the definition of the market portfolio considerably improves the performance of the CAPM in explaining the cross section of stock returns of portfolios sorted by size and beta. Recently Lettau and Ludvigson (2000b) introduced the ratio of consumption to total wealth as an important explanatory variable for the conditional CAPM and CCAPM. In their paper, total wealth contains both financial and human capital wealth.

Our work also relates to the literature on long horizon return predictability. This literature documents the forecasting power of prices scaled by dividends or earnings (Campbell and Shiller (1988a and b), Fama and French (1988a and b), Hodrick (1992), Lamont (1998)). Lettau and Ludvigson (2000a) add to this literature by showing the predictive power of the aforementioned consumption to total wealth ratio. Others have shown that various interest rate measures are also good forecasting variables (Campbell (1987), Fama and French (1989), Hodrick (1992), and Keim and Stambaugh (1986)).

Although our paper certainly provides yet more evidence on the predictability of stock returns, it differs from the ones above in its focus because we propose and test a novel economic mechanism that is able to generate time varying expected returns. The implication of our model is that the fraction of labor income to consumption should help predict stock returns through its impact on the covariance between consumption growth and returns. To lend further support to the model, we also test directly this latter implication: We use a bivariate GARCH(1,1) model to obtain a fitted time series of covariances between returns and real consumption growth and then check whether the latter can be predicted using the fraction of labor income to total consumption. Regression results strongly confirm our basic economic mechanism and give also support to the intuition that the last decade should have witnessed an increase in the covariance between returns and consumption growth due to the increase of total income produced by stock market related proceeds.

As a final test of the economic story proposed in this paper, we also look at the predictive power of the ratio of labor income to total income. Indeed, in a Lucas

(1978) pure-exchange economy we must have that consumption equals total income in equilibrium and hence the implications we obtain for the wages-to-consumption ratio should hold for the wages-to-total income ratio as well. We show that the qualitative behavior of these two series are very similar: For all subsamples, the labor income to total income ratio has a good predictive power for future returns and it is only slightly weaker for short horizons.

We should point out that our framework does not attempt to resolve other long standing issues in the field of empirical asset pricing, such as the risk-premium puzzle or the risk-free puzzle. Rather, it makes a general point about how alternative sources of income may enter the analysis to provide additional insights into the behavior of returns. For this reason we depart from the recent trend of investigating alternative preference specifications to dwell on the well known iso-elastic case, and instead concentrate on writing a flexible model that allows us to price traded assets when aggregate consumption has multiple sources. Instead of modeling the alternative sources of income (consumption) separately and then imposing some add up constraint we model the fraction of each source over the aggregate and impose that these fractions sum up to one. This apparently minor trick has deep consequences in our ability to model multiple sources of consumption in a tractable manner. Our framework will deliver closed form solutions to the pricing of both the total stock market and the individual securities.

The rest of the paper proceeds as follows. Section 2 contains the model and discussion of the assumptions. Section 3 provides the results that motivate our empirical strategy. Data description and the empirical results are reported in section 4. Section 5 concludes. All proofs, tables, and figures are in the appendix.

2 The Model

We consider a standard pure exchange economy populated by identical investors with time-separable preferences over consumption of a homogeneous good. We will assume throughout that investors are endowed with a logarithmic utility function, although this is not necessary for the results:[2] That is,

$$U(C,t) = e^{-\phi t}\log(C) \tag{1}$$

where $C(t)$ denotes investors' consumption at time t. Let there be one risky asset and one risk-free asset. The risky asset pays a continuous dividend rate $D(t)$ and its supply is normalized to 1. The riskless asset pays a continuous rate of return $r(t)$ and it is zero net supply. We also assume that investors are endowed with a stochastic labor income (rate) $w(t)$.

We assume that the dividend rate and the wage rate follow the joint processes

$$\begin{aligned} dw &= \mu_w(D,w)\,dt + \sigma_w(D,w)\,dB \\ dD &= \mu_D(D,w)\,dt + \sigma_D(D,w)dB \end{aligned}$$

where $\sigma_i(D,w)$ is a 1×2 vector for $i = D, w$, and $dB = (dB_1, dB_2)'$ is a 2×1 vector of independent Brownian motions defined on a filtered space $(\mathcal{F},\mathcal{P},\Omega)$. The specific forms of $\mu_i(D,w)$ and $\sigma_i(D,w)$ are discussed below.

Let $P(t)$ denote the price of the risky asset at time t and $N(t)$ be the number of shares bought of the risky asset at time t and let $W(t)$ be the total wealth of the investor at time t. Investors solve the following (standard) maximization problem

$$\max_{N(t),C(t)} E_0\left[\int_0^\infty U(C(s),s)ds\right]$$

subject to the dynamic budget constraint

$$dW(t) = N(t)\left(dP(t) + D(t)dt\right) + \left(W(t) - N(t)\,P(t)\right)r(t)dt + w(t)dt - C(t)dt$$

[2] In the appendix we also solve the model for the case where investors have the standard iso-elastic utility function. Since the main qualitative findings are immune to the degree of risk aversion, we chose to present the theoretical results only on this case because of its algebraic simplicity.

Definition: A *Rational Expectations Equilibrium* is given by a price function $P(t)$, an allocation process $N(t)$ and consumption $C(t)$ such that investors maximize their utility and markets clear, that is

$$N(t) = 1 \quad (2)$$

$$C(t) = D(t) + w(t) \quad (3)$$

In order to be consistent with the empirical evidence we model consumption growth as a smooth process with constant volatility:

$$\frac{dC}{C} = \mu_c dt + \sigma_c dB \quad (4)$$

where μ_c is a scalar process (possibly constant) and σ_c is a 1×2 vector of constants. Clearly (4) imposes restrictions on the process for wages and dividends. In order to meet those restrictions in a tractable, yet flexible, manner we assume a certain stochastic structure for the ratio of labor income to consumption:[3]

$$s_w = \frac{w}{C}$$

In particular we assume that $s_w(t)$ follows an autoregressive process:

$$ds_w = a\left(\overline{s}_w - s_w\right)dt + s_w\left(1 - s_w\right) v dB \quad (5)$$

where $\overline{s}_w$ is the long run share of consumption produced by labor income and a is the speed at which the share of consumption mean-reverts to $\overline{s}_w$. The form of the diffusion term, $s_w\left(1 - s_w\right)$, ensures that for every t we have $0 < s_w < 1$. v is a 1×2 vector and it can be chosen to obtain any covariance between the growth of the labor share and consumption growth. In fact

$$Cov\left(ds_w, \frac{dC}{C}\right) = s_w\left(1 - s_w\right) v \sigma_c'$$

[3] This setup is equivalent to modeling a process for the aggregate income, $I(t)$, and imposing the same stochastic structure on the shares of the different sources of income. In equilibrium, of course, $C(t) = I(t)$.

Clearly, any time s_w gets close to the boundary of zero or one, it must be the case that it has (almost) zero covariance with consumption: in the former case it is such a small fraction of total consumption that changes in this share do not affect the consumption growth itself (which is driven by dividends only). In the latter case, it is already the case that the whole consumption is generated by wages: an infinitesimal increment in the share (getting even closer to one) does not affect per se the consumption growth.

The specific assumptions about the process for consumption and for the fraction of consumption produced by labor income generate particular processes for dividends and wages through a straight application of Ito's lemma. In order to provide some economic interpretation of these processes, it is useful to go first through a few identities that holds in any model where $C = w + D$.

First, we always have

$$\frac{dC}{C} = \left(\frac{w}{C}\right)\frac{dw}{w} + \left(\frac{D}{C}\right)\frac{dD}{D} \tag{6}$$

That is, the growth rate of consumption is always a weighted average of the growth rate of wages and dividends, weighted by their relative contribution to the total consumption.

Second, since we always have

$$w = C \times \left(\frac{w}{C}\right)$$

we must also have

$$\frac{dw}{w} = \frac{dC}{C} + \left(\frac{C}{w}\right) \times \left[d\left(\frac{w}{C}\right) + Cov\left(\frac{dC}{C}, d\left(\frac{w}{C}\right)\right)\right] \tag{7}$$

Since $C/w > 0$, equation (7) shows that in order to have a growth rate of wages greater than the consumption growth we must either have an expected increase in the fraction of consumption produced by wages or a positive covariance between consumption growth and the fraction of consumption produced by wages. This is just an *accounting identity*. A similar identity holds for the growth rate of dividends.

These identities and the assumption about $\overline{s}_w$ deliver the following processes for wages and dividends:

$$dw = \mu_w(D,w)\,dt + \sigma_w(D,w)\,dB \tag{8}$$

$$dD = \mu_D(D,w)\,dt + \sigma_D(D,w)dB \tag{9}$$

with

$$\mu_w = a(\overline{w} - w) + w\mu_c + w(1 - s_w)\,v\sigma_c' \tag{10}$$

$$\mu_D = a\left(\overline{D} - D\right) + D\mu_c - Ds_w v\sigma_c' \tag{11}$$

$$\sigma_w = w(\sigma_c + (1 - s_w)\,v) \tag{12}$$

$$\sigma_D = D(\sigma_c - s_w v) \tag{13}$$

where $\overline{w} = C \times \overline{s}_w$ and $\overline{D} = C \times \overline{s}_D$.

These formulas are easy to interpret. Consider first μ_w, which has three components. The first component $a(\overline{w} - w)$ is an autoregressive part in the process for wages. The second component $w\mu_c$ makes sure that in average wages grow with the economy (consumption). The third part $w(1 - s_w)\,v\sigma_c'$ is a term involving the covariance between consumption growth and the changes in the fraction of wages as explained in equation (7).

Instead, the volatility of wages σ_w is proportional to both the volatility of consumption and the volatility of the fraction of consumption produced by wages. Clearly, when $s_w = 1$, then $\frac{1}{w}\sigma_w = \sigma_c$, that is consumption equal wages and hence they must have the same volatility. Instead, their volatilities differ whenever they are not identical.

3 Equilibrium Prices and Returns

3.1 Prices

Our modeling assumptions make possible to obtain closed form solutions for the prices of the traded assets. In our set up, as in others, labor income is a tradable asset and we abstract from any effect that market incompleteness may have on asset prices.[4] Given this assumption, the standard pricing formula applies:

$$P(t) = E_t\left[\int_t^{\infty} \frac{U_c\left(C(s), s\right)}{U_c\left(C(t), t\right)} D\left(s\right) ds\right]$$

As the next proposition shows, it turns out that we can compute the expectation directly and provide a closed form solution for the asset prices. Let Λ be the matrix defined by

$$\Lambda = \begin{pmatrix} -\lambda_{12} & \lambda_{12} \\ \lambda_{21} & -\lambda_{21} \end{pmatrix}$$

where $\lambda_{21} = a \times \overline{s}_w$ and $\lambda_{12} = a - \lambda_{21}$ and where a and $\overline{s}_w$ are the parameters defining the s_w process in equation (5).

Proposition 1 *The stock market price is*

$$P = B_{21} w + B_{22} D \tag{14}$$

where the coefficients B_{ij} are the ij elements of the matrix B given by

$$B = \left(\phi \mathbf{I} - \Lambda'\right)^{-1} \tag{15}$$

The price of stock is not only proportional to dividends, as it is generally true in the usual Lucas economy, but it has also a loading on the current wage rate. The reason why labor income w appears in the pricing function is because it helps predicting future dividends. To give a numerical example, a simple calibration of the model entails that[5]

$$B = \begin{pmatrix} 34.63 & 22.40 \\ 5.36 & 17.59 \end{pmatrix} \tag{16}$$

[4] See Campbell (1996) for example. For a lucid defense of this assumption see Jagannathan and Wang (1996, page 13).

[5] The parameters for the calibration were estimated from data and are as follows: $\phi = .025$, $a = 0.0586$, $\overline{s}_w = 0.80$, $v = (0.1063, 0)$, $\mu_c = 0.0315$; $\sigma_c = (-0.0004, 0.0164)$. All figures below are plotted using these numerical values.

and hence a pricing equation

$$P(t) = 5.36 \times w(t) + 17.59 \times D(t) \tag{17}$$

Hence, even in our model dividends have most of the weight. However, even in the pathological case where current dividends are zero, the price of the asset is not zero (although it is very small) because investors forecast future dividends to be positive.

We should point out that because of the log-utility assumptions the linear pricing formula (14) holds under any assumptions about the growth rate of total consumption. The appendix obtains a similar pricing formula for higher coefficient of risk aversion by making a mild assumption about μ_c. Since the log-utility case is sufficient to yield the intuition of the empirical implications, the results about a higher coefficient of risk aversion are left to the appendix to avoid the introduction of further notation here.

3.2 Stock Returns

We now obtain the process for stock returns. Let us denote the excess return of the asset by

$$dR = \frac{dP + Ddt}{P} - rdt$$

where r is the equilibrium instantaneous return on the risk free bond. We then have the following:

Proposition 2 *(a) The excess stock return is given by*

$$dR = \mu_R dt + \sigma_R dB \tag{18}$$

where

$$\mu_R(s_w) = \sigma_c^2 + \frac{\overline{B} s_w (1 - s_w) v \sigma_c'}{\overline{B} s_w + B_{22}} \tag{19}$$

$$\sigma_R(s_w) = \sigma_c + \frac{\overline{B} s_w (1 - s_w) v}{\overline{B} s_w + B_{22}} \tag{20}$$

and

$$\overline{B} = B_{21} - B_{22}$$

(b) The riskless rate of return is given by:

$$r = \phi + \mu_c - \sigma_c^2$$

Proposition 2 shows that the instantaneous expected return depends non-linearly on the fraction of consumption produced by labor income $s_w = w/C$. Its functional form shows that expected returns are equal to σ_c^2 both when $s_w = 0$ and when $s_w = 1$. To understand this result, notice that when $s_w = 0$, then $C = D$ and hence we are in the usual Lucas (1978) economy with no other endowment than the risky asset. The results that in the log utility case $\mu_R = \sigma_c^2$ is indeed standard. More puzzling at first is the result that $\mu_R = \sigma_c^2$ when $s_w = 1$. In this case $C = w$ and $D = 0$. However, recall that even though $D = 0$ we still have $P(t) = B_{21}w(t)$ because investors can expect that dividends will turn positive in the future. Hence, there is a perfect correlation between returns and consumption (recall again $C = w$) yielding the result.

What happens in the more realistic case where $s_w \in (0, 1)$? Indeed, from Figure 1 we can gauge that $s_w \in (0.75, 0.9)$. We can see that the denominator of the second term of μ_R in expression (19) is always positive, hence its behavior depends on $\overline{B}s_w(1 - s_w)v\sigma_c'$. Intuitively, the price of stock depends mainly on the dividend level, and hence we must have $\overline{B} = B_{21} - B_{22} < 0$ (see equation (16) and (17)). This implies that the behavior of expected returns depends on sign($v\sigma_c'$) =sign[$Cov(ds_w, dC/C)$]. If $v\sigma_c' < 0$, then we have that $\mu_R > \sigma_c^2$ with the additional implication that expected returns are decreasing in s_w for s_w sufficiently high. The opposite is true when $v\sigma_c' > 0$. Data on s_w and consumption growth suggest a negative relationship between ds_w and dC/C. Hence, intuitively as s_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. Figures 2 and 3 plot $\mu_R(s_w)$ and $\sigma_R(s_w)$ for the estimated values.

Interestingly, from Figure 3 we see that the volatility of returns is higher than both the volatility of consumption (= .0176) and the (model) volatility of ds_w which ranges between 0.01 (for high s_w) and 0.02 (for low s_w). The reason is that we can rewrite the price function as

$$P(t) = C(t)\left(\overline{B}s_w(t) + B_{22}\right)$$

Since $\overline{B} < 0$, negative innovations in consumption have a very high effect on prices because they are negatively correlated with s_w, which, by moving up, depress the price of stock even further. Hence, this combined effect tend to generate a high volatility of returns compared to consumption.

We end this section by noticing that in this model the dividend price ratio proxies for s_w. In fact, we can rewrite

$$\frac{D}{P} = \frac{1 - s_w}{\overline{B}_{21}s_w + B_{22}} \tag{21}$$

This is a nonlinear decreasing function of s_w. As it can be seen from Figure 1, this negative relationship held quite well until the middle of the 80's, when then both s_w and D/P declined together up to 1999. Indeed, a regression (not reported) of $\log(D/P)$ onto s_w produces negative coefficients, although they are strongly significant only for the sample 1946 - 1986. In any event, this result also implies that the dividend price ratio should be a predictor of future returns.

3.3 The Predictability of Long-Term Stock Returns

The preceding subsection shows that *instantaneous* returns can be expressed as a function of both the dividend price ratio and the wages to consumption ratio. Proposition 4 below shows that this representation carries over to long term returns.

Let us denote the expected excess return over the period t to $t+\tau$ by

$$R^{e}(t,t+\tau)=E_{t}\left[\frac{P(t+\tau)+\int_{t}^{t+\tau}e^{\int_{s}^{\tau}r(u)du}D(s)\,ds}{P(t)}\right]-\frac{1}{B(t,\tau)},$$

where $B(t,\tau)$ is the price of a zero coupon bond maturing at time $t+\tau$. We can easily solve for $E_t[P(t+\tau)]$ and $E_t[D(s)]$. However, since μ_c could be assumed stochastic, also the real interest rate $r(s)$ would be and hence solving for $E\left[e^{\int_s^\tau r(u)du}D(s)\right]$ presents some difficulties. We assume in this section that the growth rate of consumption μ_c is constant. This assumption can be easily justified on empirical grounds, given the extremely small autocorrelation of consumption growth. Hence, this implies that $r(u)=r$. Under this assumption, we obtain the following:

Proposition 3 *Let μ_c be constant. Then the long term expected excess returns are given by*

$$R^{e}(t,t+\tau)=F(s_{w},\tau)=\frac{\overline{B}_{P}(\tau)\,s_{w}(t)+B_{P2}(\tau)}{\overline{B}_{21}s_{w}(t)+B_{22}}-e^{r\tau}$$

and $\overline{B}_P(\tau)$ and $B_{P2}(\tau)$ are known functions of time to maturity τ specified in the appendix.

Once again, the relationship between long-term expected returns $R^e(t,t+\tau)=F(s_w,\tau)$ and s_w depends on the whether ds_w has positive or negative correlation with consumption growth. As mentioned in the previous section, data show ds_w and dC/C are negatively correlated, generating a negative relation between expected returns and s_w at short horizons. For the calibrated model, this relationship carries over to long horizons. Indeed, Figure 4 plots the instantaneous, 1,2,3, 4 year ahead expected returns $F(s_w,\tau)$ and they all are negatively sloped. Since D/P is also decreasing in s_w, the model also predicts a (non-linear) positive relationship between future returns and dividend-price ratios.

3.4 Returns to Labor

In the present model labor income is a fully tradable asset and we can mechanically compute the expected discounted value of future wages as

$$\begin{aligned} P_w(t) &= E_t\left[\int_0^\infty \frac{U'(u,C(u))}{U'(t,C(t))} w(u)du\right] && (22) \\ &= B_{11}w(t) + B_{12}D(t) && (23) \end{aligned}$$

Again, this quantity depends on both dividends and wages. Hence, we can also compute the excess returns to labor as

$$dR_w = \frac{dP_w + wdt}{P_w} - rdt.$$

As in proposition 2 we can provide a closed form expression for the returns on human capital.

Proposition 4 *Excess return on human capital follow the process*

$$dR_w = \mu_{R,w}dt + \sigma_{R,w}dW$$

where

$$\begin{aligned} \mu_{R,w}(s_w) &= \sigma_c^2 + \frac{\overline{B}_w s_w (1-s_w) v\sigma_c'}{\overline{B}_w s_w + B_{12}} \\ \sigma_{R,w}(s_w) &= \sigma_c + \frac{\overline{B}_w s_w (1-s_w) v}{\overline{B}_w s_w + B_{12}} \end{aligned}$$

and

$$\overline{B}_w = B_{11} - B_{12}$$

The implication of this proposition is that a measure of expected return to labor should also be able to predict future returns. In fact, $\mu_{R,w}(s_w)$ is increasing with s_w (for $v\sigma_c' < 0$) for s_w sufficiently high. This implies that high returns to labor should be associated with high s_w which in turn predict low future expected stock returns. The use of some proxy for return to labor income is indeed the approach taken in a number of recent articles, such as Campbell (1996) and Lettau and Ludvigson (2000a). The above set up formalizes the relationship between these quantities in a rather standard general equilibrium set up.

3.5 The Cross-Section of Stock Returns

The model presented in the previous pages can be easily extended to any number of assets. In order to do so, one has to model the joint process for the dividend to consumption ratio for all the securities. Specifically, define

$$s_1 = \frac{w}{C} \text{ and } s_i = \frac{D_i}{C} \text{ for } i = 2, ..., n$$

and let s_i follow the process

$$ds_i = [s\Lambda]_i \, dt + s_i \sigma_i (s) \, dB \tag{24}$$

where

$$\sigma_i (s) = \left(\nu_i - \sum_{j=1}^{n} s_j \nu_j \right)$$

and where Λ is a $n \times n$ matrix whose rows sum up to 0, ν_i are $1 \times n$ vectors and $B(t) = (B_1(t), .., B_n(t))'$ are n independent Brownian motions. We can see that this is a generalization of the model for s_w presented in section 2 where we set $n = 2$. In this latter case, we see immediately that taking $s_1 = s_w$ and $s_2 = (1 - s_w)$ we have

$$\begin{aligned} ds_w &= (s_w(-\lambda_{12}) + (1 - s_w)\lambda_{21})\, dt + s_w (\nu_1 - s_w \nu_1 - (1 - s_w)\nu_2)\, dB \\ &= (\lambda_{12} + \lambda_{21}) \left(\frac{\lambda_{21}}{\lambda_{12} + \lambda_{21}} - s_w \right) dt + s_w (1 - s_w)(\nu_1 - \nu_2)\, dB \end{aligned}$$

By defining

$$a = \lambda_{12} + \lambda_{21} \; ; \; \overline{s}_w = \frac{\lambda_{21}}{\lambda_{12} + \lambda_{21}} \; ; \; v = \nu_1 - \nu_2$$

we obtain equation (5) for s_w. In other words, (24) is the continuous time analog of a standard multivariate autoregressive process for the fractions s_i, with some restrictions on Λ and $\sigma_i(s)$ to make sure that $\sum_{i=1}^{n} s_i(t) = 1$ for all t.

The generalization to proposition 1 and 2 is the following

Proposition 5 *(a) The stock price of each asset is*

$$P_i(t) = B_{i1} w(t) + \sum_{j=2}^{n} B_{ij} D_j \tag{25}$$

where the coefficients B_{ij} are the ij elements of the matrix B given by

$$B = (\phi \mathbf{I} - \Lambda')^{-1} \tag{26}$$

(b) The excess return on asset i is given by

$$dR_i = \mu_i(s)\, dt + \sigma_i(s)\, dB$$

where

$$\mu_i(s) = Cov_t\left(dR_i, \frac{dC}{C}\right) = \sigma_i(s)\, \sigma_c' \tag{27}$$

and

$$\sigma_i(s) = \sigma_c + \frac{\sum_{j=1}^{n} B_{ij} s_j \left(\nu_j - \sum_{k=1}^{n} s_k \nu_k\right)}{\sum_{j=1}^{n} B_{ij} s_j} \tag{28}$$

Equations (27) and (28) show that asset i instantaneous expected excess return is a non-linear function of s_1, ..., s_n. The influence of each $s_j(t)$ on $E_t[dR_i]$ depends on the values of B_{ij} and $\nu_j \sigma_c'$. To gauge the sizes of B_{ij}'s and $\nu_j \sigma_c'$ in a full calibrated model requires the use of large amount of data and it is outside the scope of this section. We only point out that for each asset i, only $s_1 = s_w$ and s_i are likely to play an important role in the cross section. The former, s_w, because it is the largest share of consumption among all of the s_j (being equal by itself to about 75% of total consumption) and hence its movement over time is likely to affect all the returns, as a common factor. The latter, s_i, because it is likely to have the largest loading B_{ii} in the pricing equation.

We can obtain a "beta" relationship for our expected returns by using known results in asset pricing (see e.g. Duffie (1996, pg. 229-230))

$$E_t[dR_i] = \beta_i(s)\, E_t[dR^*] \tag{29}$$

where

$$\beta_i(s) = \frac{\sigma_{R,i}(s)\,\sigma'_{R^*}}{\sigma'_{R^*}\sigma'_{R^*}} \tag{30}$$

and dR^* is a return process of a portfolio perfectly correlated with the consumption process. We now show that the market portfolio augmented by human capital has this property. As in the previous section, let $P_w(t)$ denote the value of human capital and let $w(t) = D_1(t)$ for notational simplicity. Then, the value of total wealth is at time t is:

$$P_{TW}(t) = P_w(t) + \sum_{i=2}^{n} P_i(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij} D_j(t) = B\sum_{j=1}^{n} D_j(t) = BC(t) \tag{31}$$

where the middle equality stems from the fact that $\sum_{i=1}^{n} B_{ij} = B = 1/\phi$ for all j (this is a property of the log case). Consider now the return on this portfolio, which includes both financial and human capital:

$$dR_{TW} = \frac{dP_{TW} + \left(w + \sum_{i=2}^{n} D_i\right)dt}{P_{TW}} - rdt$$

From (31) it follows

$$\begin{aligned} dR_{TW} &= \frac{dC}{C} + (\phi - r)\,dt \\ &= \mu_{R,TW}dt + \sigma_{R,TW}dB \end{aligned} \tag{32}$$

with

$$\begin{aligned} E\left[dR_{TW}\right] &= \mu_c dt + (\phi - r)\,dt \\ &= \sigma_c\sigma'_c dt \\ \sigma_{R,TW} &= \sigma_c \end{aligned}$$

Hence, this portfolio satisfies the requirement for the "beta" representation because its return is perfectly correlated with the consumption process. Let finally $P_M = \sum_{i=2}^{n} P_i$ be the portfolio of financial assets. Let R_M and R_w be the excess returns to P_M and P_w, respectively. Then,

$$\begin{aligned}
dR_{TW} &= \frac{dP_w + wdt}{P_w + P_M} + \frac{dP_M + Ddt}{P_w + P_M} - rdt \\
&= \frac{P_w}{P_w + P_M} dR_w + \frac{P_M}{P_w + P_M} dR_M \\
&= \left(\sum_{i=1}^{n} s_i \frac{B_{1i}}{B} \right) dR_w + \left(1 - \sum_{i=1}^{n} s_i \frac{B_{1i}}{B} \right) dR_M
\end{aligned}$$

Now, assuming that all B_{1i} are about of the same order of magnitude, the fact that $s_w = s_1$ is by itself the largest component of all by several orders of magnitude, we can make the approximation[6]

$$dR_{TW} \approx s_w \frac{B_{11}}{B} dR_w + \left(1 - s_w \frac{B_{11}}{B} \right) dR_M$$

Plugging this last expression in (29) and recalling that $R^* = R_{TW}$, we obtain a representation of the conditional CAPM

$$\begin{aligned}
E_t \left[dR_i \right] &= \beta_i \left(s \right) E_t \left[dR_{TW} \right] \\
&= \beta_i \left(s \right) E_t \left[dR_M \right] + \widetilde{\beta}_i \left(s \right) E_t \left[s_w dR_w \right] \\
&\quad + \widetilde{\beta}_i \left(s \right) E_t \left[s_w dR_M \right]
\end{aligned} \tag{33}$$

where $\widetilde{\beta} \left(s \right) = \beta \left(s \right) B_{11} / B$.

Our model also delivers testable implications for the conditional Consumption CAPM (CCAPM). Substituting (32) in (29) and recalling that $\sigma_{R_{TW}} = \sigma_c$, we find:

$$E_t \left[dR_i \right] = \beta_i^c \left(s \right) \left(E_t \left[\frac{dC}{C} \right] + \left(\phi - r \right) dt \right) \tag{34}$$

$$= \beta_i^c \left(s \right) \left(\phi - r \right) dt + \beta_i^c \left(s \right) E_t \left[\frac{dC}{C} \right] \tag{35}$$

where

$$\beta_i^c \left(s \right) = \frac{\sigma_i \left(s \right) \sigma_c'}{\sigma_c' \sigma_c'}$$

[6] This approximation has to be compared with the one obtained by Campbell (1996, equation (10)) though empirical implementation in his work requires a formula with constant coefficients.

Taking rdt in (35) on the left hand side, we finally obtain:

$$E_t\left[dR_i^G\right]=\beta_i^c(s)\,\phi dt+\beta_i^c(s)\,E_t\left[\frac{dC}{C}\right] \tag{36}$$

where $dR_i^G=(dP_i+D_idt)/P_i$ is the return to asset i (recall that dR_i was the *excess* return).

Before concluding the section, it is worth emphasizing again the intuition behind the beta representation of expected returns in equations (29) and (30): The idea is that when the fraction of total income produced by wages decreases, expected returns for the whole market must increase, as we discussed in the previous sections. This implies that expected returns have to increase also for all the individual securities (in average). However, there is no guarantee that the risk premium increases equally across individual securities: This is because each security may have a different sensitivity to w/C, as it can be seen from the price equation (25). As an extreme example, suppose that for an asset i we have $\lambda_{ji}=0$ for j (that is, s_i is not affected by any other s_j for $j\neq i$), then one obtains $B_{ij}=0$ for all $j\neq i$. In this case, its risk premium will be different from the one of an otherwise identical security that has some $\lambda_{1i}\neq 0$. Another way of seeing this is to consider the process for dividends that it is implied by (24). This is given by

$$dD_i=\mu_{D,i}(D)\,dt+D_i\sigma_{D,i}(D)\,dB \tag{37}$$

where

$$\begin{aligned}\mu_{D,i}(D)&=D_i\mu_c+[D\Lambda]_i+\sigma_i(s)\sigma_c'\\ \sigma_{D,i}(D)&=\sigma_c+\sigma_i(s)\end{aligned}$$

These formulas are analogous to μ_D and σ_D obtained in equations (11) and (13) in section 2 and they can be interpreted in the same fashion: The first term in the drift shows a common component, so that each dividend grows with the economy. The second component show the "influence" that each dividend has on each other through the matrix

Λ. This is a sort of a multivariate autoregressive component in the dividend drift and it can be interpreted as the contribution of each dividend j (and wages, for $j = 1$) to the "production" of dividend i.[7] The last component is a covariance term that must enter into the specification, according to the general formula (7) which was developed for wage growth but which can be interpreted for dividend growth as well. As it can be seen in $\mu_{D,i}$, the current wage $w(t) = D_1(t)$ may have a different impact on the dividends of individual securities i's depending on λ_{1i}. As pointed out above, this leads to different sensitivities of prices to wages and hence different changes in risk-premia as w/C changes over time.

4 Empirical Results

4.1 Data description

The financial data we use is the usual one. 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 both consumption and labor income we use data from the National Income and Product Accounts (NIPA). Following the literature, we define consumption as non-durable plus services excluding shoes and clothing. The argument behind this idea is that the theory applies to the flow of consumption and it should not include additions or replacements to the stock of durable goods. As it is also traditional we assume that total consumption is proportional to consumption of non durables plus services and we choose the constant of proportionality, λ, to be the long term ratio between total consumption

[7] In current work, Santos and Veronesi (in progress), we explore this interpretation of our model in greater depth. In particular, we model the different sources of consumption as different sectors of the economy to which resources have to be allocated for production.

and consumption of non durables plus services. This ratio is estimated to be $\lambda = 1.15$.

As for the definition of labor income it is constructed by adding to wages and salaries, transfer payments plus other labor income and substracting personal contributions to social insurance and taxes.[8]

For the cross sectional tests we use the 25 portfolios as sorted by Fama and French (1992, 1993). These portfolios, as it is well known by now, are formed by intersecting 5 portfolios sorted by size, with other five portfolios sorted by book-to-market. We convert the returns to quarterly data producing a time series covering 1963 to 1999.

4.2 Forecasting regressions

In this section we test the forecasting power of the ratio of labor income to consumption. Table II reports the main empirical result of this paper, namely the estimates of the set of regressions

$$r_{t,t+K} = \alpha_1 + \beta_1(K)s_{w,t} + \varepsilon_{t+k} \quad (38)$$
$$r_{t,t+K} = \alpha_2 + \beta_2(K)\log\left(\frac{D_t}{P_t}\right) + \varepsilon_{t+k} \quad (39)$$
$$r_{t,t+K} = \alpha_3 + \beta_3(K)\log\left(\frac{D_t}{P_t}\right) + \beta_4(K)s_{w,t} + \varepsilon_{t+k} \quad (40)$$
$$r_{t,t+K} = \alpha_5 + \beta_5(K)\left(s_w \times \log\left(\frac{D_t}{P_t}\right)\right) + \varepsilon_{t+k} \quad (41)$$

where $r_{t,t+K}$ is the cumulative log return over K periods. For each regression, Tables II-IV report the point estimate of the included explanatory variable, the Newey-West corrected $t-$statistic for the null hypothesis that the coefficients are zero, and the adjusted R^2.

We start discussing our results for the whole sample period at our disposal, 1946-1999. The first row of Table II-A shows that the ratio of labor income to consumption

[8] We follow Lettau and Ludvigson (2000 a, footnote 15) closely in our definition of labor income and consumption. As of the writing of this paper, data from 1929 to 1945 was still being revised by the Bureau of Economic Analysis. For an update visit, *http://www.bea.doc.gov.* See also footnote 10.

s_w is statistically significant at any forecasting horizons between one quarter and four year ahead. The sign of the regression coefficient is negative, giving support to the view expressed in the previous section that innovations in consumption growth and in the wages-to-consumption ratio are negatively correlated. The explanatory power is also high, ranging from 6.1% for the one year regression to 34.6% for the four year regression. These numbers have to be compared to the analogous estimates for the (log) dividend-price ratio: In this case, the dividend price ratio is never significant at any horizons and the predictive power is rather low, ranging from 3.9% at the one year horizon to 20.7% at the four year horizon. As we will see below, this poor performance of the dividend-price ratio to forecast future returns is partly due to the inclusion of the period 1995-1999 in the data sample. Indeed, during this period the dividend price ratio was extremely low and yet returns have been record high.

Recall that, as shown in equation (21), the dividend price ratio is a non linear function of s_w. Including the dividend price ratio then may proxy for any non linearities in the relation between returns and s_w. When both ratios enter linearly (regression (40)) both regressors are significant and the predictive power ranges from an adjusted R^2 of 16.2% at the one year horizon to an R^2 of over 61% at the four year horizon. Interestingly, a similar result obtains when we only use the interaction factor $s_w \times \log(D/P)$ in the regression (regression (41)). In this case again the coefficient is highly significant and the R^2 ranges from 13.4% at the one year horizon to over 50% at the four year horizon.

As shown by Figure 1, the first two years of our sample, 1946-7, showed a remarkable drop in the ratio of labor income to consumption. The special circumstances of those years may account for that event. We reran the predictability regression excluding those eight initial data points and we report the results in Table II-B. The short term predictability of our variable improves slightly but overall the estimates are very similar. Below we perform another direct test of our theory and confirm the outlier nature of those initial two years.

Researchers like Lettau and Ludvigson (2000a) concentrate on the period starting in 1952. In order to check the validity of our empirical findings, Tables III and IV report results for two subsamples of the data starting in 1952.[9] The first is the standard sample 1952:01-1994:04, that will enable us to compare our results to others in the literature, such as the ones in Campbell, Lo and MacKinlay (1997, page 269, Table 7.1). Indeed the second line in Table III shows that during this period the dividend price ratio was doing extremely well in predicting future returns. Our results using quarterly data are comparable to the ones reported in Campbell et al. (1997) who instead used monthly data. For instance, for the one year and four year horizons they obtain R^2's of 18.8% and 41.7% respectively, matching our results almost to the point. The use of quarterly data then does not seem to be producing any particular bias in the results.

As can be seen in the first line of Table III, the ratio of labor income to consumption is now only a significant predictor at horizons of two years or more. The R^2 is 3% for the one year regression but rises to 22.6% for the four year regression. Interestingly, however, when both the dividend price ratio and the wages-to-consumption ratio are included – either in linear fashion as in regression (40) or in a nonlinear one as in regression (41) – the performance of the predictability regression improves considerably, with R^2 ranging between 25.1% for the one year regression to above 62.7% for the four year regression.

We can compare the results in Table III with those in Table IV, where the same exercise is carried out for the sample period 1952:01-1999:04. We include these results to show the dramatic decrease in the predictive power of the dividend price ratio due to the impressive stock market surge of the 90's (see e.g. Cochrane (1997) for a similar point) and to show how the labor income-to-consumption ratio still works well. Indeed,

[9] We point out that the main result of this paper (that the labor income to consumption ratio forecast future returns) accidentally underwent an out-of-sample test. Until April 26th, 2000 data on compensation of employees and other series from the Bureau of Economic Analysis were only available for the period 1959:01-1999:04. The BEA news release on April 26th, 2000 also included the revised data for the 1946:01 - 1958:04 period. Our result held well also when the first period was included.

we see that over this sample period the dividend price ratio is never significant at any horizon and the R^2 of the predictive regressions does not go above 7% at any horizon. Instead, the wages-to-consumption ratio is performing quite well: The coefficients are statistically significant at all horizons and its predictive power ranges between 7.4% at the one year horizon to 35.4% at the four year horizon. As before, using both ratios greatly improves upon the predictability regression: using regression (40) for example we obtain R^2 ranging from 19.6% at the one year to above 56% at the four year horizon.

Figures 5 and 6 emphasize this point further by plotting the time series of wages-to-consumption ratio, dividend price ratio and the four year cumulative stock return. As it can be seen in Figure 5, the dividend-price ratio and the cumulative four year return started moving in opposite directions at the end of the 80's. However, from Figure 6 we see that the negative relation between the cumulative four year return and the ratio of labor income to consumption held well even in the last part of the sample. We should notice that this behavior of cumulative returns and wages-to-consumption ratio is consistent with the model presented in section 2: A lower wages-to-consumption ratio should predict higher returns because of the increase in the equilibrium covariance between returns and consumption growth itself.

The negative relationship between the wages-to-consumption ratio and the dividend-price ratio does no longer hold when multiple assets are traded. Indeed it is easy to prove that when there are multiple stocks the dividend price ratio of the market portfolio depends on the entire distribution of dividend payments across the market (unless investors have log utility.) We explore this issue in greater depth in current work (Santos amd Veronesi (in progress)).

4.3 Cross sectional Regressions

In this section we test the implications of the model for the cross-section of stock returns, which was developed in section 3.5. Recall that the model has predictions for both the conditional CAPM and the conditional Consumption CAPM. Following Jagannathan and Wang (1996) and more recently, Lettau and Ludvgison (2000b), we test our conditional versions of these asset pricing models by "conditioning down," that is, by obtaining modified versions of their unconditional counterparts.

4.3.1 Cross-sectional Implications

The starting point of our cross sectional tests of the conditional CAPM is expression (33). Let's define $\gamma_{M,t} = E_t[dR_M]$, $\gamma_{sM,t} = E_t[s_w dR_M]$, $\gamma_{w,t} = E_t[s_w dR_w]$, $\overline{\beta}_i = E[\beta_i(s)]$, and $\overline{\widetilde{\beta}}_i = E\left[\widetilde{\beta}_i(s)\right]$. Applying the unconditional expectations operator we obtain,

$$\begin{aligned} E[dR_i] &= \overline{\beta}_i E[dR_M] + \overline{\widetilde{\beta}}_i E[s_w dR_w] + \overline{\widetilde{\beta}}_i E[s_w dR_M] \\ &\quad + cov\left(\beta_i(s), \gamma_{M,t}\right) + cov\left(\widetilde{\beta}_i(s), \gamma_{w,t}\right) \\ &\quad + cov\left(\widetilde{\beta}_i(s), \gamma_{sM,t}\right) \end{aligned} \tag{42}$$

The covariance terms distinguish the conditional version of the CAPM from its unconditional counterpart. Its interpretation is the usual one. For instance, in the case of $cov\left(\beta_i(s), \gamma_{M,t}\right)$, portfolios whose betas are high when the market premium is high will have a larger premium than the one predicted by the standard unconditional CAPM.

Direct tests of expression (42) are typically prevented by the fact that we do not observe $\beta_i(s)$, $\gamma_{M,t}$, $\widetilde{\beta}_i(s)$,and $\gamma_{w,t}$, unless they are obtained from an economic model as it is the case in this paper. As already mentioned $\gamma_{M,t}$, $\gamma_{sM,t}$ and $\gamma_{w,t}$ are mainly functions of s_w and for tractability, we approximate them by linear functions of this variable. Also, we argued in section 3.5 that $\beta_i(s)$ is mainly a function of s_w and s_i. To simplify, we approximate this function also by a linear function of these variables, i.e. we set $\beta_i(s) = a_i + b_i s_w + c_i s_i$. This yields simple expressions for (42) that can be directly tested. For instance, if we let $\gamma_{M,t} = d_i + e_i s_w$, the first covariance term becomes

$$cov\left(\beta_i(s), \gamma_{M,t}\right) = b_i e_i Var(s_w) + c_i e_i cov(s_i, s_w)$$

Notice now that with many assets, s_i is small so that we can well approximate

$$cov\left(\beta_i\left(s\right),\gamma_{M,t}\right)\approx A_i E\left(s_w^2\right)$$

for some A_i, once we notice that the constant in the variance will not play an important role. Similar calculations yields similar expressions for the other covariance terms in (42), so that overall we can write

$$E\left[dR_i\right]\approx\overline{\beta}_i E\left[dR_M\right]+\overline{\overline{\beta}}_i E\left[s_w dR_w\right]+\overline{\overline{\beta}}_i E\left[s_w dR_M\right]+\alpha_i E\left(s_w^2\right). \tag{43}$$

Expression (43) is the object of our empirical tests. Since we find that s_w^2 and s_w are extremely highly correlated, we will perform the tests using $E\left[s_w\right]$ rather than $E\left[s_w^2\right]$: The test results are almost dentical but the former lends itself better to comparisons with other work.

The starting expression for our tests of the conditional Consumption CAPM is (36). As with the conditional CAPM we obtain testable implications by "conditioning down." Let $\gamma_{c,t}=E_t\left[\frac{dC}{C}\right]$. Assuming again that $\beta_i^c\left(s\right)=a_i+b_i s_w+c_i s_i$ and applying the unconditional expectations operator we obtain,

$$E\left[dR_i^G\right]=a_i+b_i E\left[s_w\right]+c_i E\left[s_i\right]+a_i E\left[\frac{dC}{C}\right]+b_i E\left[s_w\frac{dC}{C}\right]+c_i E\left[s_i\frac{dC}{C}\right]$$

If s_i is small, we can set the third and the last term equal to zero to find

$$E\left[dR_i^G\right]=\phi a+\phi b_i E\left[s_w\right]+a_i E\left[\frac{dC}{C}\right]+b_i E\left[s_w\frac{dC}{C}\right], \tag{44}$$

and it is this last expression that we test in the empirical section. Notice that this relationship holds for simple returns and not excess returns. However, we find negligible differences in the parameter estimates when we use excess returns as opposed to simple returns (with the only obvious exception of the constant), and hence we report the results using excess returns instead.

4.3.2 Empirical Results

The 25 Fama-French portfolios have become the test ground for any model of the cross section and it is to this set that we restrict our investigations. The value premium, the fact that firms with high book to market ratios have higher average returns than firms with low book to markets, has been conclusively documented by Fama and French (1992) in the US data set. Indeed the "value premium is pervasive" in the markets across the world (see Fama and French (1998)).[10] As we will see shortly, conditioning on s_w results in considerable improvements of the CAPM and CCAPM over the unconditional versions. In order to compare the different specifications we report the R^2 in the different regressions, but the reader should bear in mind that we have only 25 (cross-sectional) observations.

A direct test of (43) requires an estimate of the return to human capital, which is not observable. Our model yields a precise method to compute those returns, by using the formula for the value of human capital given in (22) and (23). Specifically, we compute the return to human capital by

$$R^G_{w,t} = \frac{P_{w,t} - P_{w,t-1} + w_t}{P_{w,t-1}} \tag{45}$$

$$= \frac{s_{w,t} - s_{w,t-1}}{s_{w,t-1} + \kappa_1} + \frac{s_{w,t}}{\kappa_2 \times s_{w,t-1} + B_{12}} \tag{46}$$

where $\kappa_1 = B_{12}/\kappa_2$ and $\kappa_2 = B_{11} - B_{12}$.[11]

Table V reports the results for our test of the conditional CAPM as given by expression (43). The first line shows the dismal performance of the unconditional CAPM.

[10] The presence of the value premium on international data presented by Fama and French (1998) is particularly relevant as they are out of sample tests of findings that were originally made for the US and Japanese data sets. For this reason it is hard to argue that the value premium is the result of data mining. In fact the US value premium is not particularly high when compared to that of other countries in their international sample (see Table III of Fama and French (1998), page 1980.)

[11] We use the parametes obtained in section 3.1. It turns out that the choice of κ_1 and κ_2 does not affect the cross sectional results below (if anything, the one reported is the lower bound).

The beta on the value weighted return is not statistically significant, enters with the wrong sign, and the R^2 is a mere 1%, that is, the unconditional CAPM only explains 1% of the cross section of stock returns. When conditioning on the share of labor income to consumption both the coefficients on the market and s_w are significant, and the adjusted R^2 substantially improves to a more respectable 33%. The performance of the model improves when we condition on the share and the scaled market returns. In order to avoid multicollinearity in the time series regression that may complicate the inference we orthogonalize the scaled variable regressing it on market returns. The adjusted R^2 is now 38% and all variables are significant. Line 5 includes all the variables predicted by our theory, where labor-income growth is used as a measure of human capital returns (see e.g. Jagannathan and Wang (1996)). In this case, all the variables except the scaled returns to human capital are significant but there are no improvement in the (adjusted) R^2. The very last line reports the same regression as in line 5 with the only difference that excess returns to human capital are computed using formula (46). In this case, all the variables are significant and the fit improves to an (adjusted) R^2 of 43%.

That conditioning judiciously can improve the performance of the CAPM has been noted by many. For instance Jagannathan and Wang (1996) test a conditional version of the CAPM, where the variable used to condition was an interest rate variable that predicts future business conditions. When comparing the unconditional versus the conditional version the R^2 goes from 1% to approximately 30%. Jagannathan and Wang (1996) go on to provide a better definition of the market portfolio that includes labor income growth as a proxy for returns to human capital (see equation (30) in page 25 of Jagannathan and Wang (1996)). This further improves the performance of the conditional CAPM to 55.21%. However these authors investigate the performance of the conditional CAPM on the 100 portfolios sorted by size and β and do not attempt to explain the value premium. The comparison between their results and those presented in this paper should be then made with some caution.

Lines 3 and 4 of Table V can also be compared with lines 4, 5, and 6 in Table 2 of Lettau and Ludvigson (2000b). For instance, when testing the conditional CAPM, including their conditioning variable and the scaled market also yields an (adjusted) R^2-square of 21% (see line 4 of their Table 2.) Again their results improve to an impressive 71% when a measure of human capital returns is included in the regression (see line 6.)

Table VI reports tests for the unconditional and conditional CCAPM as given by expression (44). The unconditional CCAPM regression (line 1) has consumption growth as a non significant regressor though it explains 1% of the cross section of expected returns. Adding the share of labor income to consumption improves the R^2 only marginally. The conditional CCAPM where just the scaled market measure and the market itself are included performs much better. As shown in line 3 of Table VI, both variables are significant and the R^2 is now a 18.64%. When we also include the wages-to-consumption ratio the R^2 jumps to 30.7% and all variables, except consumption growth, are significant.

Figure 7 gives a quick visual impression of the considerable improvement that follows when we condition both the CAPM and the CCAPM by plotting realized average returns versus fitted returns.

5 Robustness

The results reported in the previous sections show that a fully exogenous economic variable - i.e. the labor income to consumption ratio - can predict the time series of future returns and partly explain the cross-section of returns. In this section we perform various tests to better understand the economics behind the results.

5.1 A Direct Test of the Economic Model

The economic theory that motivated our empirical investigation entails that the covariance between returns and consumption growth moves over time due to the change in the ratio between labor income and total consumption. In this section we test whether such a co-movement actually exists. We use again the quarterly data on consumption and deflate them using the CPI series to obtain a time series of real consumption growth. In order to test whether the ratio between labor income and consumption predict the level of covariance between returns and real consumption growth, we need an estimate of the time series of the latter. We fit the following bi-variate GARCH(1,1) model to data from 1946-1999:

$$\begin{pmatrix} R_t \\ \Delta c_t \end{pmatrix} = \mu + \varepsilon_t$$

where μ is a 2×1 vector of constants[12] and $\varepsilon_t \sim \mathcal{N}(0, H_t)$ with

$$H_t = C + \Delta H_{t-1} \Delta' + A \varepsilon_{t-1} \varepsilon'_{t-1} A'$$

To ensure that H_t is positive semi-definite for all t we assume $C = PP'$ where P is lower triangular. In addition, we also assume that Δ and A are diagonal matrices to limit the number of parameters.

Table VII reports the results of the Bi-variate Garch Estimates and Figure 8 plots the time series of the fitted covariance between returns and real consumption growth (dotted line). We also plot the *negative* of the labor-to-consumption ratio (solid line), rescaled to match the time series of the covariance between consumption growth and returns. If we abstract from the first two years of data (1946-1947), where a big initial outlier likely affected the measure of the covariance in the following four-to-six quarters, the plot shows a interesting association between the level of the fitted covariance between returns and consumption growth and the negative of the wages-to-consumption ratio. This is confirmed in Table VIII, where we run the following regressions

$$COV_{t+1} = \alpha_1 + \alpha_2 s_{w,t} + \alpha_3 COV_t + \varepsilon_{t+1}$$

[12] Of course, expected returns μ are not constant in this model. We make this assumption to estimate the covariance between returns and consumption in a systematic way. To be fully consistent, we should assume that μ depends on the covariance itself.

where $COV_t = \left[\widehat{H}_t\right]_{12}$, i.e the fitted covariance between consumption growth and returns, $s_{w,t} = w_t/C_t$ is the wages-to-consumption ratio. As can be seen in the table we also study specifications where α_3 is set equal to zero. Confirming the intuition developed in Figure 8, the use of the whole sample 1946-1999 results in a non significant coefficient α_2 for the wages-to-consumption ratio, due to the large outlier at the very beginning of the sample. If we eliminate the first few observations from the regression to allow for the effect of the large initial outlier to fade, we indeed find a strong negative relation between the wages-to-consumption ratio and the covariance between returns and consumption growth. This relationship holds true even when we investigate the other two relevant subsamples, 1952-1999, and 1952-1994.

5.2 Labor Income, Total Income and Consumption

The results in the previous section rely on the macro variable "labor income over consumption ratio." Strictly speaking, the model proposed in this paper also implies that total consumption equals total income in equilibrium. As a consequence, the same relationships that hold for the labor income to consumption ratio should also hold for the labor income to total income ratio. In this section we take on this point and show that there is a strict relationship between labor-income to consumption ratio and the labor-income to total income ratio. For coherence with the labor income series (which is net of taxes), we take the total *disposable* income as a measure of total income.[13] Figure 9 plots the two series for our sample. As it can be seen, the general qualitative time series behavior of these series are very similar, with the labor income to total income being slightly smoother than the other one. In order to test for predictability, Table IX shows

[13]It is worth reminding the relationship between total income and disposable income in the national accounts (NIPA tables). We have Total Personal Income equals Compensation of Employees plus Proprietors Income plus Rental Income plus Personal Dividend Income plus Personal Interest Income plus Transfer payments to persons *less* Personal Contribution to Social Insurance. Finally, Disposable Income equal Total Income less Personal Taxes.

the predictive regressions using the labor income to total income series as predictors over the various samples. As it can be seen, the labor income to disposable income ratio has a very good predictive power for future returns in all subsamples, although the results are not so strong for the one quarter ahead predictive regression..

5.3 Spurious Regressions

Most series used in this paper are close-to-unit root series and hence there is a justified concern that the results about the time-series predictability may be due to spurious regressions (see e.g. Torous and Yan (1999) for a recent work about spurious predictive regressions.[14]) To summarize the problem, suppose that we have the pair of series

$$y_{t+1} = a + bx_t + u_{t+1} \tag{47}$$

$$x_{t+1} = \alpha + \phi x_t + v_{t+1} \tag{48}$$

where $y_t = \log(R_t)$ is the log return and where $\varepsilon_t = (u_t, v_t)'$ is a martingale difference sequence such that $E[\varepsilon_t \varepsilon_t' | \varepsilon_{t-1}, \varepsilon_{t-2} ...] = \Sigma$ and such that u_t and v_t have only contemporaneous correlation $\rho = Corr(u_t, v_t)$. Following Richardson and Stock (1989), Tourus and Yan (1999) study the asymptotic distribution of the least square estimator obtained by regressing the long-term return $y_{t+1}(K) = \sum_{i=1}^{K} y_{t+i}$ onto the predictive variable x_t under the null hypothesis that there is no relation between y_t and x_t but through the contemporaneous correlation of the series. That is, the null-hypothesis is $\mathcal{H}_0 : b = 0$. They show that indeed the for given K, the estimated $\beta(K)$ is consistent as $T \to \infty$, but its Newey-West adjusted t-statistic $t_{\beta(K)}$ has a non-standard distribution which depends on the correlation ρ between u_t and v_t. If $\rho = 0$, then $t_{\beta(K)}$ has indeed a standard normal distribution. They also show that if the number of non-overlapping intervals does not grow to infinity with T, that is if $T/K \to c$, constant, then $\beta(K)$ is no longer consistent and $t_{\beta(K)}$ has a non-standard distribution that depends on both ρ and c.

[14] Aside from the original article by Phillips (1986), the reader is referred to Chap. 18.3 in Hamilton (1994) for a lucid exposition of the spurious regression problem and for other references.

In order to check the robustness of our results, we take the suggestions contained in Torous and Yan (1999) and obtain more robust confidence intervals for the coefficient $\beta(K)$ in our predictive regressions by means of Monte Carlo Simulations. More specifically, we perform the following exercise: Let x_t in equation (48) be any of the regressors used in the forecasting regressions, i.e. the labor-income to consumption ratio $s_{w,t}$, the (log) dividend price ratio $\log(D_t/P_t)$, both of them or the interaction term $s_{w,t}\log(D_t/P_t)$. For each of them, we first compute the parameters α and ϕ in equation (48) from a time-series regression and the matrix $\Sigma = E[\varepsilon_t \varepsilon_t']$ to take into account the correlation between u_t and v_t (recall that if the correlation $\rho = 0$, then $\beta(K)$ is indeed consistent as $T \to \infty$ and $t_{\beta(K)}$ is indeed distributed as a standard normal distribution. In this case all the results in the previous section hold.) Given our sample size $T = 216$ for the period 1946-1999, we simulate 10,000 paths of the system (47)-(48) under the null hypothesis that $b = 0$. For each sample, we compute the predictive regressions as in equations (38)-(41) and obtain a distribution for $\beta(K)$. This is tabulated in Table X. We repeated the experiment using both the estimations of the relevant parameters of (47)-(48) and the sample sizes corresponding to the periods 1952-1999 and 1952-1994 and obtained extremely similar cutoff values for $\beta(K)$, which we do not report for brevity.

For the sample period 1946-1999 the estimated coefficients reported in Table II-A are extremely close to being statistically significant at the 10% level for the one quarter, one year, and two year regression. For example, the one quarter ahead predictive regression yields a coefficient $\beta(1) = -.308$ while from the simulations, the 90% confidence interval is $[-0.318, 0.302]$. For the three and four year regression both coefficients are statistically significant at the 10% level. For instance for the four year regression the estimated coefficient is given by $\beta(16) = -4.902$ while the 90% confidence interval is $[-4.514, 4.394]$.

In contrast the dividend price ratio is never significant at the 10%. This is partly due to the fact that the correlation between returns and dividend price ratio

$\rho = Corr(u_t, v_t)$ is much higher in this case (around $-.96$) than in the case where the predictive variable is the labor income to consumption ratio (in which case $\rho \approx -.02$). As explained above, a higher correlation ρ makes the distribution of $t_{\beta(K)}$ even more "non-normal". For example, using the dividend price ratio for the one quarter and four year regression we obtain $\beta(1) = .022$ and $\beta(K) = .475$ respectively while robust confidence intervals are $[-0.005, 0.0388]$ and $[-0.09, 0.551]$ respectively.

Looking at the regression where both the share and the log dividend price ratio enter linearly the estimated coefficient on s_w is significant at all horizons. For instance, for the one quarter predictability regression the estimated value of the coefficient is $\beta(1) = -.467$, whereas the 90% confidence interval is given by $[-.416, .365]$. Once again, the dividend price ratio is never significant at this level of statistical significance.

The results are even stronger when the sample is restricted to the period 1948-1999 (see Table II-B).[15] The estimated coefficients for the univariate regression with s_w are all significant at the 10% level with the exception of the one year ahead regression, which is, in any case, extremely close to the cut-off value. All of them are statistically significant in the regression where the dividend price ratio enters linearly.

Overall, this section confirms from a statistical point of view that the predictive regression results obtained in section 4.2 are unlikely to be spurious, and it lends further credence to the hypothesis of an economic relation linking returns on the aggregate market portfolio and the share of labor income to consumption.

6 Conclusion

We propose a simple general equilibrium model to show that changes in expected returns may be generated by changes in the relative importance of various sources of income. In our model, total income is funded by dividends and labor income that grow stochas-

[15] The table of the simulated distribution for the sample period 1948-1999 is extremely close to Table X when the simulation are done with eight observations less. We do not report it for brevity.

tically over time. We show that equilibrium expected returns change as the fraction of total income funded by labor income fluctuates over time, because the latter affect the conditional covariance between equilibrium returns and consumption growth. We then obtain a new and simple testable implication, namely, that the ratio of labor income to consumption should help predicts 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 that are larger than those generated when using the dividend price ratio as a single explanatory variable. Notice that differently from the dividend price ratio, our variable is a pure macroeconomic variable that is constructed with no reference to financial variables and it is in this sense that we provide a powerful test of our theory.

We derive and test a version of the conditional CAPM and the conditional consumption CAPM, where the variable we use to condition is the aforementioned ratio. Tests are conducted on the justly celebrated 25 Fama-French portfolios. For both the CAPM and the CCAPM we indeed find that when we use our labor-income to consumption ratio as conditioning variable, the fit in the cross section of returns is substantially improved when compared to the unconditional version of the models. We interpret this finding as lending additional support to our conditioning variable.

In addition to the previous tests on the time series and cross section predictability, we also provide a direct test of the main mechanism that produces changes in expected returns in our model, namely, that the covariance between returns and consumption growth should be predicted by the ratio of labor income to total consumption. We find a strong support also for this relationship in the data and notice that the covariance between returns and consumption growth increased steadily from the beginning of the 80's.

In conclusion, we find a substantial support for the economic model proposed in this paper, that is, that the time variation in the relative importance of the various sources of income have an important effect on the required expected return. We should point out that our model is also fully consistent with a recent body of literature documenting that cyclical variation in the consumption to wealth ratio affect future expected returns. Although our model is casted in the log-utility case for simplicity, which implies a constant consumption to wealth ratio, the same results of this paper can be obtained with higher degrees of risk aversions as shown in the appendix. In this case, one can show that the consumption to wealth ratio is a non-linear but decreasing function of the labor income to consumption ratio. Hence, this model predicts that the consumption to wealth ratio should also be a predictor of future returns, a finding that is well documented in Lettau and Ludvigson (2000a).

7 References

Barberis, Nick, Ming Huang and Tano Santos (2000), "Prospect Theory and Asset Prices," forthcoming, *Quarterly Journal of Economics*.

Campbell, John Y. (1987), "Stock Returns and the Term Structure," *Journal of Financial Economics,* 18, 373-399.

Campbell, John Y. (1996), "Understanding Risk and Return," *Journal of Political Economy*, 104, 298-345.

Campbell, John Y. (1999), "Asset Prices, Consumption, and the Business Cycle," in Handbook of Macroeconomics, John B. Taylor and Michael Woodford editors.

Campbell, John Y. (2000), "Asset Pricing at the Millennium." *Forthcoming Journal of Finance*, August.

Campbell, John Y. and John H. Cochrane (1999), "By Force of Habit," *Journal of Political Economy.*

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay (1997), The Econometrics of Financial Markets. Princeton University Press, Princeton, New Jersey.

Campbell, John Y. and Robert J. Shiller (1988 a), "The Dividend Pprice Ration and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies,* 1, 195-227.

Campbell, John Y. and Robert J. Shiller (1988 b), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661-676.

Cochrane, John H. (1996), "A Cross Sectional Test of an Investment Based Asset Pricing Model," *Journal of Political Economy,* 104 (3), 572-621.

Cochrane, John H. (1997), "Where is the Market Going? Uncertain Facts and Novel Theories," *Economic Perspectives*, Federal Reserve of Chicago, November-December..

Cochrane, John H. (2000), "New Facts in Finance," *Economic Perspectives. Federal Reserve Bank of Chicago.*

Cochrane, John H. (2000), Asset Pricing, Manuscript, University of Chicago.

Duffie, Darrell (1996), Dynamic Asset Pricing Theory, Princeton University Press, Pricenton, New Jersey.

Fama, Eugene F. and James D. MacBeth (1972), "Risk, Return, and Equilibrium: Empirical Tests," Journal of Political Economy, 81, 607-636.

Fama, Eugene F. and G. William Schwert (1977), "Human Capital and Capital Market Equilibrium," *Journal of Financial Economics,* 4, 115-146.

Fama, Eugene F. and Kenneth R. French (1988 a), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics,* 22, 3-27.

Fama, Eugene F. and Kenneth R. French (1988 b), "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy,* 96, 246-273.

Fama, Eugene F. and Kenneth R. French (1989), "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics,* 25, 23-49.

Fama, Eugene F. and Kenneth R. French (1992), "The Cross Section of Expected Returns," *Journal of Finance,* 47, 427-465.

Fama, Eugene F. and Kenneth R. French (1993), "Common Risk Factors in the Returns of Stocks and Bonds," *Journal of Financial Economics,* 33, 3-56.

Ferson, Wayne E. and Campbell R. Harvey (1999), "Conditioning Variables and the Cross Section of Stock Returns," *Journal of Finance*, 54, 1325-1361.

Jagannathan, R. and Zhenyu Wang (1996), "The Conditional CAPM and the Cross Section of Expected Returns," *Journal of Finance,* 51, 3-54.

Hamilton, James D. (1994), Time Series Analysis, Princeton University Press, Princeton, New Jersey.

Hodrick, Robert (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies,* 5, 357-386.

Keim, Donald B. and Robert F. Stambaugh (1986), "Predicting Stock Returns in the Stock and Bond Markets," *Journal of Financial Economics,* 17, 357-390.

Lamont, Owen (1998), "Earnings and Expected Returns," *Journal of Finance*, 53, 1563-1587.

Lettau, Martin and Sydney Ludvigson (2000a), "Consumption, Aggregate Wealth and Expected Stock Returns," working paper, Federal Reserve of New York.

Lettau, Martin and Sydney Ludvigson (2000b), "Resurrecting the (C)CAPM: A Cross-Sectional Test of Linear Factor Models When Risk Premia are Time-Varying," working paper, Federal Reserve of New York.

Lipster, R. S. and A. N. Shiryayev (1977), Statistics of Random Process Vol. 1, Springer Verlag, Berlin, Heidelberg, Nnew York.

Lucas, Robert E. (1978), "Asset Prices in an Exchange Economy," *Econometrica* 46, 1429-1445.

Mayers, David, (1972) "Nonmarketable Assets and Capital Market Equilibrium under Uncertainty," in M. Jensen ed.: *Studies in the Theory of Capital Markets.* Praeger, New York, 223-248.

Phillips, Peter C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics,* 33, 311-340.

Richadson, Matthew and James Stock (1989), "Drawing inferences from statistics based on multiyear asset returns," *Journal of Financial Economics,* 25, 323-348.

Santos, Tano and Pietro Veronesi (in progress), "General Equilibrium and the Cross Section of Stock Returns," *Manuscript,* University of Chicago.

Shiller, Robert J. (1981), "Do prices move too much to be explained by subsequent changes in dividends?," *American Economic Review,* 71, 421-436.

Timmerman, Allan (1993), "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 108, 1135-1145.

Timmerman, Allan (1996), "Excess Volatility and Predictability of Stock Returns in Autoregressive Dividend Models with Learning," *Review of Economic Studies*, 1996, 523-557.

Torous, Walter and Shu Yan (1999), "Predictive Regressions Revisited," working paper, UCLA.

Veronesi, Pietro (2000), "How Does Information Quality Affect Stock Returns?" *Journal of Finance*, 55, 807-837.

Vissing-Jørgensen, Annette (1999), "Limited stock market participation and the equity premium puzzle" Working paper, University of Chicago.

Appendix

Existence of the share process

To prove the existence of the share process we introduce some extra notation that will also be needed to prove some claims about stock prices and stock returns. Consider n shares $s_1, ..., s_n$ such that $\sum_{i=1}^{n} s_i = 1$. Suppose they jointly follow the process

$$\begin{aligned} ds_i &= [s\Lambda]_i \, dt + s_i \sigma_i (s) \, dB \\ &= \left(\sum_{j=1}^{n} s_j \lambda_{ji} \right) dt + s_i \sigma_i (s) \, dB \end{aligned} \tag{49}$$

for all $i = 1, 2, \ldots, n$ where Λ is a $n \times n$ matrix given by

$$\Lambda = \begin{pmatrix} -\sum_{j \neq 1} \lambda_{1j} & \lambda_{11} & ... & \lambda_{1n} \\ \lambda_{21} & -\sum_{j \neq 2} \lambda_{2j} & ... & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & ... & -\sum_{j \neq n} \lambda_{nj} \end{pmatrix} \tag{50}$$

and $\sigma_i(s)$ is a $1 \times n$ vector given by

$$\sigma_i(s) = \nu_i - \sum_{j=1}^{n} s_j \nu_j \tag{51}$$

We prove here that a solution to the system (49) exists and that for all t, we have

$$\begin{aligned} s_i (t) &> 0 \\ \sum_{i=1}^{n} s(t)_i &= 1 \end{aligned}$$

Specialization to $n = 2$ covers the case discussed in the theory section of the paper.

To prove the claim above we can appeal to an analogous result on optimal filtering obtained by Liptser and Shyriaev (1977) where the $s_i(t)$ are interpreted as the probability that an unobserved variable $z(t)$ is in state z_i at time t. Specifically, let a $n \times 1$ dimensional state variable $z(t)$ follow a continuous time, regime shift model with infinitesimal transition matrix

Λ and state i being given by z_i (an $n \times 1$ vector). Let the observation process be a n vector $X(t)$ following the process

$$dX = z(t) + \Sigma dB^*$$

where Σ is $n \times n$. Also, let

$$s_i(t) = \Pr\left(z(t) = z_i | \left(X(\tau)\right)_{\tau=0}^{t}\right)$$

then a simple extension to the multivariate case of Theorem 9.1, page 333, in Liptser and Shyriaev (1977) shows that

$$ds_i = [s\Lambda]_i \, dt + s_i \left(z_i - \sum_{j=1}^{n} s_j z_j\right)' \left(\Sigma\Sigma'\right)^{-\frac{1}{2}} dB$$

where

$$dB = \left(\Sigma\Sigma'\right)^{-\frac{1}{2}} \left\{dX - E_t\left(dX\right)\right\}$$

is a multidimensional Wiener process with respect to the filtration generated by $\left(X(\tau)\right)_{\tau=0}^{t}$. Defining

$$v_i = z_i' \times \left(\Sigma\Sigma'\right)^{-\frac{1}{2}}$$

we obtain the system (49). Existence of a unique solution and the other properties are proved in the above set-up in Theorem 9.1, page 333, and Lemma 9.3, page 342 in Liptser and Shiryayev (1977). In our model, we only use the structure of the process (49) and not the original interpretation.

Proof of propositions 1

In this appendix we prove that a linear pricing function obtains for both the case where investors have log-utility and the case where they have the iso-elastic utility function

$$U\left(C, t\right) = e^{-\phi t} \frac{C^{1-\gamma}}{1-\gamma}$$

Specifically, we prove the following:

Proposition A1 *Let either of the following assumptions hold:*

1. *Investors have log-utility:* $U(t, C) = e^{-\phi t} \log\left(C\right)$.

2. *Investors have power utility* ($\gamma \neq 1$) *but consumption growth is* $\mu_c = \overline{\mu}_c + s_w\theta_w + (1-s_w)\,\theta_D$, *where* $\overline{\mu}_c$ *is constant and* θ_w *and* θ_D *are two constant satisfying* $\theta_w = v\sigma_c' + \theta_D$;

Then the stock market price is

$$P = B_{21}w + B_{22}D$$

where the coefficients B_{ij} *are the* ij *elements of the matrix* B *given by*

$$B = \left(\phi\mathbf{I} - \Lambda'\right)^{-1}$$

in case (a); and

$$B = \left(\widehat{\Theta} - \Lambda'\right)^{-1}$$

with $\widehat{\Theta} = diag\left(\widehat{\theta}_w, \widehat{\theta}_D\right)$ *with* $\widehat{\theta}_i = \phi + (\gamma-1)\left(\overline{\mu}_c + \theta_i\right) - \frac{1}{2}\gamma\left(\gamma-1\right)\sigma_c^2$ *in case (b), respectively.*

To prove this proposition it is convenient to work with the general set of processes (49), that with the specification

$$ds_i = [s\Lambda]_i\,dt + s_i\sigma_i\left(s\right)dB \tag{52}$$

with

$$\sigma_i\left(s\right) = \left(\nu_i - \sum_{j=1}^{n} s_j\nu_j\right)$$

where in this paper we have $n = 2$. In this case, we see immediately that taking $s_1 = s_w$ and $s_2 = (1-s_w)$ we have

$$\begin{aligned} ds_w &= \left(s_w(-\lambda_{12}) + \left(1-s_w\right)\lambda_{21}\right)dt + s_w\left(\nu_1 - s_w\nu_1 - \left(1-s_w\right)\nu_2\right)dB \\ &= \left(\lambda_{12}+\lambda_{21}\right)\left(\frac{\lambda_{21}}{\lambda_{12}+\lambda_{21}} - s_w\right)dt + s_w\left(1-s_w\right)\left(\nu_1-\nu_2\right)dB \end{aligned}$$

By defining

$$k = \lambda_{12}+\lambda_{21}\ ;\ \overline{s}_w = \frac{\lambda_{21}}{\lambda_{12}+\lambda_{21}}\ ;\ v = \nu_1 - \nu_2$$

we obtain the stochastic differential equation (5). It turns out that it is simpler to use the system of SDEs (52) rather than the single SDE (5). Of course, the approaches are identical. In this case, however, we must reformulate the assumption about the growth rate of consumption. Specifically, we assume

$$\mu_c = \overline{\mu}_c + \sum_{i=1}^{2} s_i \theta_i$$

where $\theta_i = \nu_i \sigma_c'$. We see that this is the same assumption above in proposition A1(b), because we have

$$\theta_1 - \theta_2 = (\nu_1 - \nu_2)\, \sigma_c' = v \sigma_c'$$

as it was assumed.

The proof of these results rely strongly on the following lemma (which holds for any n).

Lemma A2: *For all* $i = 1, .., n$, *let* $\mu_c = \overline{\mu}_c + \sum_{j=1}^{n} s_j \theta_j$, $\widehat{\theta}_i = \phi - \beta(\overline{\mu}_c + \theta_i) - \frac{1}{2}\beta(\beta - 1)\sigma_c^2$ *and* $\widehat{\Theta} = diag\left(\widehat{\theta}_1, ..., \widehat{\theta}_n\right)$. *Then*

$$E\left[C(u)^{\beta} s_i(u)\right] = \sum_{k=1}^{n} C(t)^{\beta} s_k(t) \sum_{j=1}^{n} w_{jk}^{-1} w_{ij} e^{\omega_j (u-t)}$$

where ω_j *are the eigenvalues of* $\overline{\Lambda}' = \left(\Lambda' + \phi \mathbf{I} - \widehat{\Theta}\right)$ *and* w_{ij} *are associated eigenvectors and* $w_{ij}^{-1} = [W^{-1}]_{ij}$.

Proof of Lemma A2: Let

$$X_i(t) = C(t)^{\beta} s_i(t) \tag{53}$$

Apply Ito's lemma to find let

$$\begin{aligned}
dX_i &= \beta C^{\beta-1} s_i dC + \frac{1}{2}\beta(\beta-1)C^{\beta-2} s_i dC^2 && (54)\\
&\quad + C^{\beta} ds_i + \beta C^{\beta-1} ds_i dC && (55)\\
&= \beta C^{\beta-1} s_i C \mu_c dt + \beta C^{\beta-1} s_i C \sigma_c dW && (56)\\
&\quad + \frac{1}{2}\beta(\beta-1) C^{\beta-2} s_i C^2 \sigma_c^2 dt && (57)\\
&\quad + C^{\beta} [s\Lambda]_i dt + C^{\beta} s_i \sigma_i(s) dW + \beta C^{\beta-1} s_i \sigma_i(s) \sigma_c' C && (58)\\
&= \left\{\beta X_i \mu_c + \frac{1}{2}\beta(\beta-1) X_i \sigma_c^2 + [X\Lambda]_i + \beta X_i \sigma_i(s) \sigma_c'\right\} dt && (59)\\
&\quad + X_i \{\beta \sigma_c + \sigma_i(s)\} dW && (60)
\end{aligned}$$

This stochastic differential equation is non-linear in the drift, due to the covariance term $\beta X_i \sigma_i(s)\sigma_c' = \beta X_i \left(\theta_i - \sum_{j=0}^n s_v \theta_j\right)$. However, since we assumed $\mu_c = \overline{\mu}_c + \sum_{j=1}^n s_j \theta_j$, we then obtain

$$\begin{aligned} dX_i &= \left\{\beta X_i \overline{\mu}_c + \frac{1}{2}\beta(\beta-1) X_i \sigma_c^2 + [X\Lambda]_i + \beta X_i \theta_i\right\} dt \\ &\quad + X_i \{\beta\sigma_c + \sigma_i(s)\} dW \\ &= \left[X\left(\Lambda + \phi \mathbf{I} - \widehat{\Theta}\right)\right]_i dt + X_i \{\beta\sigma_c + \sigma_i(s)\} dW \end{aligned}$$

Defining $\overline{\Lambda} = \left(\Lambda + \phi\mathbf{I} - \widehat{\Theta}\right)$ and $\widehat{\sigma}_i(s) = \beta\sigma_c + \sigma_i(s)$ we can rewrite

$$dX_i = \left[X\overline{\Lambda}\right]_i dt + X_i \widehat{\sigma}_i(s) dW$$

Using the vector notation, with $X = (X_1, ..., X_n)$ as a $1 \times n$ vector, we can rewrite this in its integral form as

$$X(u) = X(t) + \int_t^u X(\tau)\overline{\Lambda} d\tau + \int_t^u X(\tau)\widehat{\Sigma}(s(\tau)) dW(\tau) \tag{61}$$

which is identical to (66). Notice that also that for all i, $X_i(t) < C_i(t)$, hence all the integral below exist as long as the expected present value of future consumption can be computed (in particular, as in the proof of proposition 1, we can apply Fubini Theorem and hence invert the integrals). Using the same steps, we then obtain

$$\begin{aligned} \widetilde{X}_i(u) &= E\left(X(u)|X(t)\right) = \sum_{j=1}^n \sum_{k=1}^n w_{jk}^{-1} X_k(t) w_{ij} e^{\omega_j (u-t)} \\ &= \sum_{k=0}^n X_k(t) \sum_{j=0}^n w_{jk}^{-1} w_{ij} e^{\omega_j (u-t)} \end{aligned}$$

where ω_j are the eigenvalues of $\overline{\Lambda}'$ and w_{ij} are associated eigenvectors and $w_{ij}^{-1} = [W^{-1}]_{ij}$. This concludes the proof of Lemma A2.■

Proof of Proposition A1: (a) For the case of logarithmic utility, we can solve for prices using the usual pricing equation

$$P(t) = E_t\left[\int_t^\infty \frac{U_C(C(u),u)}{U_C(C(t),t)} D(u) du\right] \tag{62}$$

$$= \quad E_t \left[\int_t^{\infty} e^{-\phi(u-t)} \frac{C(t)}{C(u)} \left(s_D(u) C(u) \right) du \right] \tag{63}$$

$$= \quad C(t) E_t \left[\int_t^{\infty} e^{-\phi(u-t)} s_D(u) du \right] \tag{64}$$

$$= \quad C(t) \int_t^{\infty} e^{-\phi(u-t)} E_t \left[s_D(u) \right] du \tag{65}$$

where the last equality stems from Fubini's theorem, whose assumptions are satisfied because the process $s_D(u) \in [0,1]$ for all u (and hence all the integrals exist). We solve this expectations under the general assumption that there is a vector of shares $s(t) = (s_1(t), s_2(t))$ following the general process (52).

Notice that the assumptions made on the process s allow us to obtain $E_t[s(u)]$ in closed form. Indeed, we can rewrite in vector form

$$s(u) = s(t) + \int_t^u s(\tau) \Lambda d\tau + \int_t^u s(\tau) \Sigma(s(\tau)) dB(\tau) \tag{66}$$

where $\Sigma(s(\tau))$ is some (bounded) $n \times n$ function of $s(\tau)$. Let $\widetilde{s}(u) = E_t(s(u))$. Since $\Sigma(s)$ is bounded,

$$E_t \left[\int_t^u s(\tau) \Sigma(s(\tau)) dW(\tau) \right] = 0$$

Hence, taking expectations on both sides with respect to time t we obtain

$$\widetilde{s}(u) = s(t) + \int_t^u \widetilde{s}(\tau) \Lambda d\tau$$

or

$$\frac{d\widetilde{s}}{d\tau} = \widetilde{s} \Lambda$$

This is a linear system of differential equations with initial condition $\widetilde{s}(0) = s(t)$. If Λ' admits real and distinct eigenvalues, its general solution is then given by

$$\widetilde{s}_i(u) = \sum_{j=1}^{2} k_j w_{ij} e^{\omega_j (u-t)}$$

where ω_j are the eigenvalues of Λ' and w_{ij} are associated eigenvectors. From the initial condition $\widetilde{s}(0) = s(t)$ we obtain that $s(t) = W \times \kappa$ where $W = [w_1, .., w_n]$ is the matrix whose columns are the eigenvectors of Λ'. Hence, $\kappa = W^{-1} \times s(t)$or

$$\kappa_j = \sum_k w_{jk}^{-1} s_k(t)$$

which implies

$$\begin{aligned}\widetilde{s}_i(u) &= \sum_{j=1}^{2}\sum_{k=1}^{2} w_{jk}^{-1} s_k\left(t\right) w_{ij} e^{\omega_j(u-t)} \\ &= \sum_{k=1}^{2} s_k\left(t\right) \sum_{j=1}^{2} w_{jk}^{-1} w_{ij} e^{\omega_j(u-t)}\end{aligned}$$

where $w_{ij}^{-1} = [W^{-1}]_{ij}$. Hence,

$$\begin{aligned}\frac{P_i}{C} &= \sum_{k=1}^{2} s_k(t) \sum_{j=1}^{2} w_{jk}^{-1} w_{ij} \int_t^{\infty} e^{-(\phi-\omega_j)(u-t)} du \\ &= \sum_{k=1}^{n} s_k(t) \left(\sum_{j=1}^{2} w_{jk}^{-1} w_{ij} \frac{1}{\phi - \omega_j} \right)\end{aligned} \tag{67}$$

Hence for $i = 1, 2$ and $k = 1, 2$

$$B_{ik} = \left(\sum_{j=0}^{2} \frac{w_{ij} w_{jk}^{-1}}{\phi - \omega_j} \right)$$

Notice that if we define the *square* matrix

$$B = \left(\mathbf{I}\phi - \Lambda'\right)^{-1}$$

then

$$B_{ik} = e_i \left(\mathbf{I}\phi - \Lambda'\right)^{-1} e_j'$$

In fact, notice that

$$\begin{aligned}B\left(\mathbf{I}\phi - \Lambda'\right) &= \mathbf{I} \\ BW\left(\mathbf{I}\phi - \Omega\right) W^{-1} &= \mathbf{I}\end{aligned}$$

where Ω is the diagonal matrix with the eigenvalues ω_j of Λ' on the principal diagonal. Hence:

$$B = W\left(\mathbf{I}\phi - \Omega\right)^{-1} W^{-1}$$

where of course $D = (\mathbf{I}\phi - \Omega)^{-1}$ is diagonal. Therefore, the ik element of B can be written as

$$\begin{aligned}B_{ik} &= \sum_j \sum_{\ell} w_{ij} D_{j\ell} w_{\ell k}^{-1} \\ &= \sum_{j=1}^{2} \frac{w_{ij} w_{jk}^{-1}}{\phi - \omega_j}\end{aligned}$$

(b): Consider again the usual pricing equation

$$
\begin{aligned}
P_i(t) &= E_t\left[\int_t^\infty \frac{U_C(C(u),u)}{U_C(C(t),t)}D(u)du\right] \\
&= E_t\left[\int_t^\infty e^{-\phi(u-t)}\left(\frac{C(u)}{C(t)}\right)^{-\gamma}(s_2(u)C(u))\,du\right] \\
&= C(t)^\gamma E_t\left[\int_t^\infty e^{-\phi(u-t)}C(u)^{1-\gamma}s_2(u)du\right]
\end{aligned}
$$

where $i = 2$ is the index defining the share $\frac{D}{C}$, that is $s_2 = s_D$.

From (53) making $\beta = 1-\gamma$ and using Lemma A2:

$$
\begin{aligned}
P_i(t) &= C(t)^\gamma E_t\left[\int_t^\infty e^{-\phi(u-t)}C(u)^{1-\gamma}s_2(u)du\right] \\
&= C(t)^\gamma \sum_{k=1}^{2} C(t)^{1-\gamma}s_k(t)\sum_{j=1}^{2} w_{jk}^{-1}w_{2j}\int_t^\infty e^{-(\phi-\omega_j)(u-t)}du \qquad (68)\\
&= C(t)\sum_{k=1}^{2} s_k(t)\left(\sum_{j=1}^{2} w_{jk}^{-1}w_{2j}\frac{1}{\phi-\omega_j}\right) \qquad (69)\\
&= \sum_{k=1}^{2} C(t)s_k(t)\left(\sum_{j=1}^{2} w_{jk}^{-1}w_{2j}\frac{1}{\phi-\omega_j}\right) \qquad (70)\\
&= \sum_{k=1}^{2} D_k(t)\left(\sum_{j=0}^{2} w_{jk}^{-1}w_{2j}\frac{1}{\phi-\omega_j}\right) \qquad (71)
\end{aligned}
$$

We finally prove that

$$
\left(\sum_{j=1}^{2} w_{jk}^{-1}w_{2j}\frac{1}{\phi-\omega_j}\right) = B_{2k} = e_2\left(\widehat{\Theta}-\Lambda'\right)^{-1}e_k = e_2\left(\phi\mathbf{I}-\overline{\Lambda}'\right)^{-1}e_k
$$

By definition $\overline{\Lambda} = \left(\Lambda + \phi\mathbf{I} - \widehat{\Theta}\right)$ and hence $\overline{\Lambda}' = \left(\Lambda' + \phi\mathbf{I} - \widehat{\Theta}\right)$. As before, notice that

$$
\begin{aligned}
B\left(\phi\mathbf{I}-\overline{\Lambda}'\right) &= \mathbf{I} \\
BW\left(\phi\mathbf{I}-\Omega\right)W^{-1} &= \mathbf{I}
\end{aligned}
$$

where Ω is the diagonal matrix with the eigenvalues ω_j of $\overline{\Lambda}'$ on the principal diagonal. Hence:

$$
B = W\left(\mathbf{I}\phi-\Omega\right)^{-1}W^{-1}
$$

where of course $D = (\mathbf{I}\phi - \Omega)^{-1}$ is diagonal. Therefore, the ik element of B can be written as

$$B_{ik} = \sum_j \sum_\ell w_{ij} D_{j\ell} w_{\ell k}^{-1} = \sum_{j=1}^{2} \frac{w_{ij} w_{jk}^{-1}}{\phi - \omega_j}$$

This concludes the proof. ■

Since the assumptions about the growth rate of consumption are difficult to interpret (although in the calibration they have a minor role: μ_c turns out to be almost constat) we can obtain an approximate similar result by assuming a constant growth rate of consumption and that approximately $s_w (1 - s_w) v\sigma_c' \approx \overline{s}_w (1 - \overline{s}_w) v\sigma_c'$. In this case, we have the following proposition:

Proposition A2: Let $\gamma \neq 1$ and $\mu_c = \overline{\mu}_c$ be a constant. Define

$$\begin{aligned} \overline{\theta}_w &= \phi + \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 + (\overline{\mu}_c + \overline{\sigma})(\gamma - 1) \\ \overline{\theta}_D &= \phi + \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 + (\overline{\mu}_c - \overline{\sigma})(\gamma - 1) \end{aligned}$$

with $\overline{\sigma} = \overline{s}_w (1 - \overline{s}_w) v\sigma'$, $\overline{\Theta} = diag\left(\overline{\theta}_w, \overline{\theta}_D\right)$ and where $\overline{s}_w$ is the (long term) stationary distribution of the process $s_w(t)$. The price of the asset is approximately given by

$$P_i \approx B_{21} w + B_{22} D$$

where the coefficients B_{ij} are the ij elements of the matrix B given by

$$B = \left(\overline{\Theta}\mathbf{I} - \Lambda'\right)^{-1}$$

Proof of Proposition A2: This is almost identical to the proof of proposition 1, part (b): from (59)-(**??**) we see that the stochastic differential equation for X_i is non-linear in the drift, due to the covariance term $(1-\gamma)X_i\sigma_i(s)\sigma_c'$. Notice that $\sigma_i(s)\sigma_c' = \theta_i - \sum_j s_j\theta_j$ is bounded above and below by $+/-\max(\theta) - \min(\theta)$. In addition, in the long run expected value of $\sigma_i(s)\sigma_c'$ is just $\theta_i - \sum_i s_i^*\theta_i$ where $s^* = (s_1^*, ..., s_n^*)$ is the stationary distribution. We make then the following approximation:

$$(1-\gamma)X_i\sigma_i(s)\sigma_c' \approx (1-\gamma)X_i\sigma_i(s^*)\sigma_c'$$

which is constant. So, let

$$\overline{\Lambda} = \Lambda + \phi \mathbf{I} - diag\left(\overline{\theta}\right) \tag{72}$$

where $\overline{\theta}_i = \phi + \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 + \left(\overline{\mu}_c - \sigma_i(s^*)\sigma_c'\right)(\gamma - 1)$. As in the proof of Lemma A2, we can write:

$$dX_i = \left[X\overline{\Lambda}\right]_i dt + X_i \widehat{\sigma}_i(s) dW$$

and hence the same steps lead to

$$P_i = \sum_{k=1}^{2} D_k(t) \left(\sum_{j=1}^{2} w_{jk}^{-1} w_{ij} \frac{1}{\phi - \omega_j} \right)$$

where ω_j are the eigenvalues of $\overline{\Lambda}'$ and w_{ij} are associated eigenvectors and $w_{ij}^{-1} = [W^{-1}]_{ij}$. As before, we have

$$\left(\sum_{j=1}^{2} w_{jk}^{-1} w_{ij} \frac{1}{\phi - \omega_j} \right) = B_{ik} = e_i \left(\overline{\Theta} - \Lambda'\right)^{-1} e_k = e_i \left(\phi \mathbf{I} - \overline{\Lambda}'\right)^{-1} e_k$$

where $\overline{\Lambda}$ is defined in (72). This concludes the proof. ■

Proof of proposition 2: We start with point (b). The riskless rate must satisfy

$$rdt = -E_t \left[\frac{dm}{m} \right]$$

where $m(t) = U_c(t, C) = e^{-\phi t} C^{-\gamma}$ is the pricing kernel. An application of Ito's lemma yields the result

$$r = \phi + \gamma \mu_c + \frac{1}{2}\gamma(\gamma + 1)\sigma_c^2$$

(a) This proof holds for both the log utility and the iso-elastic case. In the case of log utility just set $\gamma = 1$. Rewrite the price as

$$P = C \sum_{i=1}^{2} s_i B_{2i}$$

From Ito's lemma we have

$$dP = dC \times \sum_{i=1}^{2} s_i B_{2i} + C \times \sum_{i=1}^{2} B_{2i} ds_i + \sum_{i=1}^{2} B_{2i} ds_i dC$$

$$
\begin{aligned}
&= \mu_c C \sum_{i=1}^{2} s_i B_{2i} dt + C \sum_{i=1}^{2} s_i B_{2i} \sigma_c dB + C \sum_{i=1}^{2} B_{2i} \left[s\Lambda \right]_i dt + C \sum_{i=1}^{2} B_{2i} s_i \left(\nu_i - \sum_{j=1}^{2} s_j \nu_j \right) dB \\
&\quad + C \sum_{i=1}^{2} B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^{2} s_j \nu_j \sigma_c' \right) dt \\
&= P \left(\mu_c + \frac{\sum_{i=1}^{2} B_{2i} \left[s\Lambda \right]_i}{\sum_{i=1}^{2} s_i B_{2i}} + \frac{\sum_{i=1}^{2} B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^{2} s_j \nu_j \sigma_c' \right)}{\sum_{i=1}^{2} s_i B_{2i}} \right) dt \\
&\quad + P \left(\sigma_c + \frac{\sum_{i=1}^{2} B_{2i} s_i \left(\nu_i - \sum_{j=1}^{2} s_j \nu_j \right)}{\sum_{i=1}^{2} s_i B_{2i}} \right) dB
\end{aligned}
$$

Hence, since

$$dR = \frac{dP}{P} + \frac{D}{P} dt - r dt$$

we have

$$
\begin{aligned}
\mu_R \left(s_w \right) &= \mu_c + \frac{\sum_{i=1}^{2} B_{2i} \left[s\Lambda \right]_i}{\sum_{i=1}^{2} s_i B_{2i}} + \frac{\sum_{i=1}^{2} B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^{2} s_j \nu_j \sigma_c' \right)}{\sum_{i=1}^{2} s_i B_{2i}} \\
&\quad + \frac{1}{\sum_{i=1}^{2} s_i B_{2i}} - \phi - \gamma \mu_c + \frac{1}{2} \gamma \left(\gamma + 1 \right) \sigma_c^2
\end{aligned}
$$

Notice also that from

$$B \left(\phi \mathbf{I} - \overline{\Lambda}' \right) = \mathbf{I}$$

we have

$$B \widehat{\Theta} - B \Lambda' = \mathbf{I}$$

or

$$B \Lambda' = B \widehat{\Theta} - \mathbf{I}$$

Hence, since

$$
\begin{aligned}
\sum_i B_{2i} \left[s\Lambda \right]_i &= \sum_i B_{2i} \sum_k s_k \lambda_{ki} = \sum_k s_k \sum_i B_{2i} \lambda_{ki} \\
&= \sum_k s_k e_2 B \Lambda' e_k = \sum_k s_k e_2 B \widehat{\Theta} e_k - 1 \\
&= \sum_k s_k B_{2k} \widehat{\theta}_k - 1
\end{aligned}
$$

From the definition of $\widehat{\theta}_k = \phi + (\gamma - 1)(\overline{\mu}_c + \theta_\kappa) - \frac{1}{2}\gamma(\gamma - 1)\sigma_c^2$ we notice that if $\gamma = 1$ we then have $\theta_k = \phi$ for all k. Hence, indeed we find that $\sum_i B_{2i}[s\Lambda]_i = \phi \sum_k s_k B_{2k} - 1$. This yields immediately

$$\begin{aligned}\mu_R(s) &= \sigma_c^2 + \frac{\sum_{i=1}^2 B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^2 s_j \nu_j \sigma_c'\right)}{\sum_{i=1}^2 s_i B_{2i}} \\ &= Cov(dC/C, dR)\end{aligned}$$

If $\gamma \neq 1$, we still have

$$\sum_k s_k B_{2k} \widehat{\theta}_k = \left(\phi + (\gamma - 1)\overline{\mu}_c - \frac{1}{2}\gamma(\gamma - 1)\sigma_c^2\right) \sum_k s_k B_{2k} + (\gamma - 1) \sum_k s_k B_{2k} \theta_k$$

Hence, recalling the assumption in this case that $\mu_c = \overline{\mu}_c + \sum_{i=1}^2 s_i \theta_i$ with $\theta_i = \nu_i \sigma_c'$ we obtain

$$\begin{aligned}\mu_R(s_w) &= (1 - \gamma)\mu_c + \frac{\sum_{i=1}^2 B_{2i}[s\Lambda]_i}{\sum_{i=1}^2 s_i B_{2i}} + \frac{\sum_{i=1}^2 B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^2 s_j \nu_j \sigma_c'\right)}{\sum_{i=1}^2 s_i B_{2i}} \\ &\quad + \frac{1}{\sum_{i=1}^2 s_i B_{2i}} - \phi + \frac{1}{2}\gamma(\gamma + 1)\sigma_c^2 \\ &= (1 - \gamma)\sum_{i=1}^2 s_i \theta_i + \gamma \sigma_c^2 + (\gamma - 1)\frac{\sum_k s_k B_{2k} \theta_k}{\sum_{i=1}^2 s_i B_{2i}} \\ &\quad + \frac{\sum_{i=1}^2 B_{2i} s_i \nu_i \sigma_c'}{\sum_{i=1}^2 s_i B_{2i}} - \sum_{j=1}^2 s_j \nu_j \sigma_c' \\ &= \gamma\left(\sigma_c^2 + \frac{\sum_{i=1}^2 B_{2i} s_i \left(\nu_i \sigma_c' - \sum_{j=1}^2 s_j \nu_j \sigma_c'\right)}{\sum_{i=1}^2 s_i B_{2i}}\right) \\ &= \gamma Cov(dC/C, dR)\end{aligned}$$

where we used again the assumption that $\theta_i = \nu_i \sigma_c'$. ■

Next lemma will be useful in proposition 4.

Lemma A3: *Assume $\mu_c = \overline{\mu} + \sum_{j=0}^n s_j \theta_j$. Then the real price of long term bond is given by*

$$B(t, \tau) = \sum_{k=1}^n s_k A_{Bk}(\tau)$$

where

$$A_{Bk} = e^{-\phi\tau} \sum_{i=0}^n \sum_{j=0}^n w_{jk}^{-1} w_{ij} e^{\omega_j \tau}$$

Proof of Lemma A3: The real bond price at time t for delivery of a unit of consumption good at time $t+\tau$ must satisfy

$$B(t,\tau) = E_t\left[\frac{U_c(t+\tau)}{U_c(t)}\right] = \frac{1}{C(t)^{-\gamma}} e^{-\phi\tau} E_t\left[C(t+\tau)^{-\gamma}\right]$$

Notice that we always have

$$C(t)^{-\gamma} = \sum_{i=0}^{n} s_i(t)\, C(t)^{-\gamma}$$

becase $\sum_{i=0}^{n} s_i(t) = 1$. Hence, we can evaluate the expectation by using the result in Lemma A2 by setting $\beta = -\gamma$. That is

$$E\left[s_i(t+\tau)\, C(t+\tau)^{-\gamma}\right] = \sum_{k=1}^{n} s_k(t)\, C(t)^{-\gamma} \sum_{j=1}^{n} w_{jk}^{-1} w_{ij} e^{\omega_j \tau}$$

where w_{ij} and ω_j are the eigenvectors and eigenvalues of the matrix $\overline{\Lambda}' = \left(\Lambda' + \phi \mathbf{I} - \widehat{\Theta}\right)$ with $\beta = -\gamma$. Hence

$$\begin{aligned} E_t\left[C(t+\tau)^{-\gamma}\right] &= \sum_{i=0}^{n} E_t\left[s_i(t+\tau)\, C(t+\tau)^{-\gamma}\right] \\ &= C(t)^{-\gamma} \sum_{i=0}^{n} \sum_{k=0}^{n} s_k(t) \sum_{j=0}^{n} w_{jk}^{-1} w_{ij} e^{\omega_j \tau} \end{aligned}$$

Hence

$$\begin{aligned} B(t,\tau) &= e^{-\phi\tau} \sum_{i=1}^{n} \sum_{k=1}^{n} s_k(t) \sum_{j=1}^{n} w_{jk}^{-1} w_{ij} e^{\omega_j \tau} \\ &= \sum_{k=1}^{n} s_k(t)\, A_{Bk}(\tau) \end{aligned}$$

where

$$A_{Bk} = e^{-\phi\tau} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{jk}^{-1} w_{ij} e^{\omega_j \tau}$$

This concludes the proof of Lemma A3. ■

Proof of Proposition 3: We start by computing the expected future price. We know that $P(t) = B_{21} w(t) + B_{22} D(t)$. Hence,

$$\begin{aligned} E_t\left[P(t+\tau)\right] &= B_{21} E_t\left[w(t+\tau)\right] + B_{22} E_t\left[D(t+\tau)\right] \\ &= B_{21} E_t\left[s_1(t+\tau)\, C(t+\tau)\right] + B_{22} E_t\left[s_2(t+\tau)\, C(t+\tau)\right] \end{aligned}$$

We can now use the results from Lemma A2 with $\beta = 1$ to obtain

$$
\begin{aligned}
E_t\left[s_i(u)\,C(u)\right] &= \sum_{k=1}^{2} s_k(t)\,C(t)\sum_{j=1}^{2} w_{jk}^{-1} w_{ij} e^{\omega_j(u-t)} \\
&= w(t)\sum_{j=1}^{2} w_{j1}^{-1} w_{ij} e^{\omega_j(u-t)} + D(t)\sum_{j=1}^{2} w_{j2}^{-1} w_{ij} e^{\omega_j(u-t)}
\end{aligned}
$$

where w_{ij} and ω_j are the eigenvectors and eigenvalues of the matrix $\overline{\Lambda}' = \left(\Lambda' + \phi\mathbf{I} - \widehat{\Theta}\right)$, with $\beta = 1$. Hence,

$$
\begin{aligned}
E_t\left[P(t+\tau)\right] &= B_{21}E_t\left[s_1(t+\tau)\,C(t+\tau)\right] + B_{22}E_t\left[s_2(t+\tau)\,C(t+\tau)\right] \\
&= B_{21}\left\{ w(t)\sum_{j=0}^{2} w_{j1}^{-1} w_{1j} e^{\omega_j\tau} + D(t)\sum_{j=0}^{2} w_{j2}^{-1} w_{1j} e^{\omega_j\tau} \right\} \\
&\quad + B_{22}\left\{ w(t)\sum_{j=0}^{2} w_{j1}^{-1} w_{2j} e^{\omega_j\tau} + D(t)\sum_{j=0}^{2} w_{j2}^{-1} w_{2j} e^{\omega_j\tau} \right\} \\
&= w(t)\,A_{P1}(\tau) + D(t)\,A_{P2}(\tau)
\end{aligned}
$$

where

$$
\begin{aligned}
A_{P1}(\tau) &= B_{21}\sum_{j=0}^{2} w_{j1}^{-1} w_{1j} e^{\omega_j\tau} + B_{22}\sum_{j=0}^{2} w_{j1}^{-1} w_{2j} e^{\omega_j\tau} \\
A_{P2}(\tau) &= B_{21}\sum_{j=0}^{2} w_{j2}^{-1} w_{1j} e^{\omega_j\tau} + B_{22}\sum_{j=0}^{2} w_{j2}^{-1} w_{2j} e^{\omega_j\tau}
\end{aligned}
$$

Similarly, we obtain the result for the expected future dividend

$$
E_t\left[\int_t^{t+\tau} e^{\int_s^{t+\tau} r(u)du} D(s)\,ds\right] = \int_t^{t+\tau} E_t\left[e^{\int_s^{t+\tau} r(u)du} s_2(s)\,C(s)\right] ds
$$

This is difficult to evaluate due to the stochastic nature of the interest rate when μ_c is stochastic as in proposition A1 (b). If $\gamma = 1$ and μ_c is a constant, then r is also a constant and we won't have any troubles. When $\mu_c = \overline{\mu}_c + \sum_{i=1}^{2} s_i\theta_i$, then the interest rate is stochastic. However, even in this case the interest rate is almost constant. We then solve the expectation under the assumption that $r(u) = \overline{r}$ so that $\int_s^{t+\tau} r(u)\,du = \overline{r}(t+\tau-s)$. Notice that this this holds for proposition 3 (log-utility) and proposition A2. We can take the interest rate out of the expectation and then use the results of Lemma A2 to find

$$\begin{aligned}
E_t\left[s_2\left(u\right)C\left(u\right)\right] &= \sum_{k=1}^{2} s_k\left(t\right)C\left(t\right)\sum_{j=1}^{2} w_{jk}^{-1}w_{2j}e^{\omega_j(u-t)} \\
&= w\left(t\right)\sum_{j=1}^{2} w_{j1}^{-1}w_{2j}e^{\omega_j(u-t)} + D\left(t\right)\sum_{j=1}^{2} w_{j2}^{-1}w_{2j}e^{\omega_j(u-t)}
\end{aligned}$$

where w_{ij} and ω_j are the eigenvectors and eigenvalues of the matrix $\overline{\Lambda}' = \left(\Lambda' + \phi\mathbf{I} - \widehat{\Theta}\right)$, with $\beta = 1$. Hence,

$$\begin{aligned}
E_t\left[\int_t^{t+\tau} e^{\overline{r}(t+\tau-u)}D\left(u\right)du\right] &= \int_t^{t+\tau} e^{\overline{r}(t+\tau-u)}E_t\left[s_2\left(u\right)C\left(u\right)\right]du \\
&= w\left(t\right)\sum_{j=1}^{2} w_{j1}^{-1}w_{2j}\int_t^{t+\tau} e^{\overline{r}(t+\tau-u)}e^{\omega_j(u-t)}du \\
&\quad +D\left(t\right)\sum_{j=1}^{2} w_{j2}^{-1}w_{2j}\int_t^{t+\tau} e^{\overline{r}(t+\tau-u)}e^{\omega_j(u-t)}du \\
&= w\left(t\right)\sum_{j=1}^{2} w_{j1}^{-1}w_{2j}\int_t^{t+\tau} e^{\overline{r}\tau}e^{(\omega_j-\overline{r})(u-t)}du \\
&\quad +D\left(t\right)\sum_{j=1}^{2} w_{j2}^{-1}w_{2j}\int_t^{t+\tau} e^{\overline{r}\tau}e^{(\omega_j-\overline{r})(u-t)}du \\
&= w\left(t\right)\sum_{j=0}^{2} w_{j1}^{-1}w_{2j}e^{\overline{r}\tau}\frac{e^{(\omega_j-\overline{r})\tau}-1}{\omega_j-\overline{r}} + D\left(t\right)\sum_{j=0}^{2} w_{j2}^{-1}w_{2j}e^{\overline{r}\tau}\frac{e^{(\omega_j-\overline{r})\tau}-1}{\left(\omega_j-\overline{r}\right)} \\
&= w\left(t\right)A_{D1}\left(\tau\right) + D\left(t\right)A_{D2}\left(\tau\right)
\end{aligned}$$

Hence, we finally have

$$\begin{aligned}
R^e\left(t,t+\tau\right) &= E\left[\frac{P\left(t+\tau\right)+\int_t^{t+\tau} e^{\overline{r}(t+\tau-s)}D\left(u\right)du}{P(t)}\right] - \frac{1}{B\left(t,\tau\right)} \\
&= \frac{1}{P\left(t\right)}\left\{B_{P1}\left(\tau\right)w\left(t\right)+B_{P2}\left(\tau\right)D\left(t\right)\right\} - \frac{1}{B\left(t,\tau\right)} \\
&= \frac{B_{P1}\left(\tau\right)w\left(t\right)+B_{P2}\left(\tau\right)D\left(t\right)}{B_{21}w\left(t\right)+B_{22}D\left(t\right)} - \frac{1}{B\left(t,\tau\right)} \\
&= \frac{B_{P1}\left(\tau\right)s_w\left(t\right)+B_{P2}\left(\tau\right)\left(1-s_w\left(t\right)\right)}{B_{21}s_w\left(t\right)+B_{22}\left(1-s_w\left(t\right)\right)} - \frac{1}{s_wA_{B1}\left(\tau\right)+\left(1-s_w\right)A_{B2}\left(\tau\right)}
\end{aligned}$$

where

$$B_{Pi}\left(\tau\right) = A_{Pi}\left(\tau\right) + A_{Di}\left(\tau\right)$$

By defining $\overline{B}_P(\tau) = B_{P1}(\tau) - B_{P2}(\tau)$ the proof is completed. ■

Proof of Proposition 4: This proof is analogous to the proof of proposition 3. The only difference is that $\overline{B}$ is now $\overline{B}_w = B_{11} - B_{12}$.

TABLE I

Summary Statistics: 1946:01 - 1999:04

	mean (quarterly)	st.dev. (quarterly)	1st. Autcorrelation	β OLS	Dickey Fuller
Returns	0.0172	0.0798	0.0442	0.024	-
$\log(D/P)$	-3.36	0.34	0.9746	0.9987	$-.2819$
w/C	0.8307	0.0374	0.9753	0.9876	-2.674

Correlation Matrix

	Returns	$\log(D/P)$	w/C
Returns	1	-	-
$\log(D/P)$	-0.1271	1	-
w/C	-0.1338	0.4306	1

Summary statistics of time series data. The last two columns report the value of the regression coefficient of an OLS regression on own lagged variable. The Dickey-Fuller statistic is also reported. Rejection of unit-root hypothesis at 1%, 5% and 10% level is for statistics below -13.6, -8.0 and -5.7, respectively.

TABLE II A

Forecasting Future Returns

Sample: 1946:01 - 1999:04

Forecasting Horizon

K	1	4	8	12	16
w/C	-0.308	-1.121	-2.275	-3.588	-4.902
t-stat.	(-2.140)	(-2.407)	(-3.029)	(-3.721)	(-4.028)
(adj) R^2	0.016	0.061	0.143	0.250	0.346
$log(D/P)$	0.022	0.105	0.213	0.338	0.475
t-stat.	(1.394)	(1.648)	(1.445)	(1.590)	(1.840)
(adj) R^2	0.003	0.039	0.084	0.142	0.207
w/C	-0.467	-1.660	-2.995	-4.237	-5.315
t-stat.	(-2.955)	(-2.960)	(-3.525)	(-4.371)	(-4.733)
$log(D/P)$	0.043	0.173	0.312	0.435	0.538
t-stat.	(2.800)	(3.008)	(3.150)	(3.931)	(4.229)
(adj) R^2	0.038	0.162	0.317	0.484	0.613
$w/C \times log(D/P)$	0.059	0.243	0.456	0.654	0.826
t-stat.	(2.992)	(3.083)	(3.358)	(4.769)	(6.221)
(adj) R^2	0.028	0.134	0.264	0.398	0.509

The table shows the result of the predictive regression

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

where $\mathbf{x}_t = w_t/C_t$; or $\log\left(D_t/P_t\right)$, or both; where K is the numbers of quarter ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. Number in parenthesis show Newey-West adjusted t-statistics. The sample is 1946:01-1999:04.

TABLE II B

Forecasting Future Returns

Sample: 1948:01 - 1999:04

Forecasting Horizon

K	1	4	8	12	16
w/C	-0.342	-1.211	-2.493	-3.761	-4.920
t-stat.	(-2.315)	(-2.548)	(-3.289)	(-3.674)	(-3.761)
(adj) R^2	0.020	0.071	0.165	0.266	0.334
$log(D/P)$	0.023	0.109	0.223	0.333	0.456
t-stat.	(1.425)	(1.689)	(1.486)	(1.550)	(1.736)
(adj) R^2	0.004	0.042	0.091	0.137	0.190
w/C	-0.552	-1.918	-3.485	-4.674	-5.598
t-stat.	(-3.315)	(-3.381)	(-4.500)	(-4.916)	(-4.780)
$log(D/P)$	0.050	0.198	0.359	0.469	0.557
t-stat.	(3.113)	(3.484)	(4.077)	(4.676)	(4.737)
(adj) R^2	0.051	0.196	0.385	0.529	0.617
$w/C \times log(D/P)$	0.065	0.266	0.503	0.685	0.834
t-stat.	(3.130)	(3.303)	(3.744)	(4.844)	(6.063)
(adj) R^2	0.034	0.153	0.303	0.416	0.496

The table shows the result of the predictive regression

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

where $\mathbf{x}_t = w_t/C_t$; or $\log(D_t/P_t)$, or both; where K is the numbers of quarter ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. Number in parenthesis show Newey-West adjusted t-statistics. The sample is 1948:01-1999:04.

TABLE III

Forecasting Future Returns

Sample: 1952:01 - 1994:04

Forecasting Horizon

K	1	4	8	12	16
w/C	-0.297	-1.037	-2.124	-3.131	-4.291
t-stat.	(-1.425)	(-1.459)	(-1.979)	(-2.300)	(-2.610)
(adj) R^2	0.007	0.030	0.081	0.145	0.226
$log(D/P)$	0.085	0.347	0.597	0.730	0.779
t-stat.	(3.055)	(3.779)	(3.507)	(3.458)	(3.577)
(adj) R^2	0.043	0.188	0.330	0.404	0.414
w/C	-0.430	-1.428	-2.456	-3.243	-4.128
t-stat.	(-2.143)	(-2.264)	(-3.459)	(-4.260)	(-4.122)
$log(D/P)$	0.098	0.377	0.624	0.740	0.763
t-stat.	(3.496)	(4.221)	(4.539)	(4.532)	(5.014)
(adj) R^2	0.063	0.251	0.443	0.564	0.627
$w/C \times log(D/P)$	0.116	0.438	0.733	0.888	0.966
t-stat.	(3.836)	(4.108)	(5.111)	(6.007)	(7.202)
(adj) R^2	0.067	0.255	0.451	0.566	0.616

The table shows the result of the predictive regression

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

where $\mathbf{x}_t = w_t/C_t$; or $\log(D_t/P_t)$, or both; where K is the numbers of quarter ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. Number in parenthesis show Newey-West adjusted t-statistics.

TABLE IV

Forecasting Future Returns

Sample: 1952:01 - 1999:04

Forecasting Horizon

K	1	4	8	12	16
w/C	-0.345	-1.218	-2.370	-3.542	-4.687
t-stat.	(-2.324)	(-2.575)	(-3.170)	(-3.595)	(-3.809)
(adj) R^2	0.021	0.074	0.160	0.259	0.354
$log(D/P)$	0.016	0.086	0.177	0.251	0.310
t-stat.	(0.823)	(1.117)	(0.943)	(0.894)	(0.935)
(adj) R^2	-0.002	0.018	0.042	0.056	0.066
w/C	-0.621	-2.175	-3.754	-4.887	-5.786
t-stat.	(-3.463)	(-3.433)	(-4.300)	(-4.720)	(-4.748)
$log(D/P)$	0.061	0.238	0.419	0.522	0.559
t-stat.	(2.856)	(3.059)	(3.263)	(3.475)	(3.689)
(adj) R^2	0.050	0.196	0.365	0.486	0.565
$w/C \times log(D/P)$	0.074	0.309	0.578	0.758	0.862
t-stat.	(2.683)	(2.799)	(3.093)	(3.803)	(4.637)
(adj) R^2	0.029	0.140	0.275	0.359	0.399

The table shows the result of the predictive regression

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

where $\mathbf{x}_t = w_t/C_t$; or $\log\left(D_t/P_t\right)$, or both; where K is the numbers of quarter ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. Number in parenthesis show Newey-West adjusted t-statistics.

TABLE V

CAPM Fama-MacBeth Regressions

	CONST	R_M	s_w	$s_w \times R_M$	$s_w \times \Delta \ln(w)$	$s_w \times R_w$	R^2 [$Adj\ R^2$]
1	2.63	$-.379$					1%
t-stat.	(2.73)	$(-.31)$					[−3%]
2	5.455	-3.677	4.506				39%
t-stat.	(4.12)	(-2.47)	(2.8)				[33%]
3	3.429	-1.235		.314			22%
t-stat	(3.38)	(-1.01)		(3.15)			[14%]
4	5.541	-3.730	3.877	.199			46%
t-stat	(4.14)	(-2.49)	(2.56)	(2.65)			[38%]
5	5.307	-3.445	3.033	0.258	-0.406		48%
t-stat	(4.07)	(-2.35)	(2.05)	(3.087)	(-1.57)		[37%]
6	5.244	-3.538	3.817	0.201		-0.586	53%
t-stat	(4.011)	(-2.39)	(2.511)	(2.665)		(-3.421)	[43%]

This table presents estimates of cross-sectional Fama-MacBeth Regressions using the 25 Fama-French portfolios. R_M is the excess returns of the value weighted market portfolio, s_w is the ratio of labor income to consumption, $s_w \times R_M$ denotes the residual of a regression of scaled excess return on excess return themselves, R_w denotes our measure of human capital excess return and $\Delta \ln(w)$ denotes real labor-income growth. R^2 denotes the unadjusted R-square.

TABLE VI

CCAPM Fama-MacBeth Regressions

	CONST	Δc	s_w	$s_w \times \Delta c$	R^2 $[Adj\ R^2]$
1	1.475	0.2042			5.21%
	(1.968)	(.667)			[1%]
2	2.623	-.0699	2.0238		12.2%
	(3.372)	(−0.289)	(1.763)		[4%]
3	1.7073	.5935		−.0189	25.42%
	(2.224)	(2.1524)		(−2.337)	[18.64%]
4	3.4108	.2706	2.9816	−.0225	39.4%
	(3.7867)	(1.3542)	(2.3316)	(−2.5773)	[30.7%]

This table presents estimates of cross-sectional Fama-MacBeth Regressions using the 25 Fama-French portfolios. Δc denotes consumption growth, s_w is the share of labor income to consumption, $s_w \times \Delta c$ denotes the residual of a regression of scaled consumption growth on consumption growth itself. R^2 denotes the unadjusted R-square.

Table VII

Estimates of Bivariate GARCH(1,1) Model

	P_{11}	P_{21}	P_{22}	Δ_{11}	Δ_{22}	A_{11}	A_{22}
Estimate	0.0460	0.0008	0.0015	-0.7739	-0.9257	0.2498	-0.2676
t-stat	(3.147)	(3.107)	(2.8934)	(-5.214)	(-34.175)	(2.391)	(-4.9398)

Results of the estimate of Bivariate GARCH(1,1) model

$$\begin{pmatrix} R_t \\ \Delta c_t \end{pmatrix} = \mu + \varepsilon_t$$

where μ is a 2×1 vector of constants and $\varepsilon_t \sim \mathcal{N}(0, H_t)$ with

$$H_t = C + \Delta H_{t-1} \Delta' + A \varepsilon_{t-1} \varepsilon'_{t-1} A',$$

$C = PP'$ where P is lower triangular and Δ and A are diagonal matrices.

Table VIII:

Predicting the Return - Consumption Growth Covariance

	Constant	$w(t-1)/C(t-1)$	Cov(t-1)	(Adj.) R^2
1946-1999				
Est	2.7751e-004	-2.234e-004		.010
t-stat	(1.0794)	(-0.7357)		
Est	1.1080e-004	-9.5139e-005	0.6537	0.4321
t-stat	(1.2553)	(-0.8980)	(21.5061)	
1948 to 1999				
Est	5.1440e-004	-4.9770e-004		.1025
t-stat	(4.0610)	(-3.1708)		
Est	1.8707e-004	-1.8389e-004	0.6631	0.5003
t-stat	(3.1824)	(-2.6930)	(10.43)	
1952 to 1999				
Est	5.2777e-004	-5.1502e-004		.1108
t-stat	(4.1641)	(-3.2824)		
Est	1.7223e-004	-1.7031e-004	0.6920	0.5385
t-stat	(3.1132)	(-2.6308)	(12.9816)	
1952 to 1994				
Est	5.8619e-004	-5.8358e-004		0.0871
t-stat	(3.1777)	(-2.6281)		
Est	1.9702e-004	-1.9946e-004	0.6923	0.5261
t-stat	(2.5103)	(-2.1904)	(12.6387)	

Results of the regression $COV_{t+1} = \alpha_1 + \alpha_2 s_{w,t} + \alpha_3 COV_t + \varepsilon_{t+1}$ where $COV_t = \left[\widehat{H}_t\right]_{12}$ is the fitted covariance between returns and consumption growth and $s_{w,t} = w_t/C_t$ is the labor income to consumption ratio.

TABLE IX

Forecasting Future Returns

	Forecasting Horizon				
	1	4	8	12	16
1946 to 1999					
Labor/Disposable	-0.354	-1.694	-3.827	-5.817	-8.051
t-stat.	(-1.126)	(-1.710)	(-2.309)	(-2.982)	(-3.836)
(adj) R^2	0.003	0.040	0.129	0.230	0.360
1952 to 1999					
Labor/Disposable	-0.729	-2.898	-5.855	-7.983	-9.261
t-stat.	(-1.720)	(-2.307)	(-2.903)	(-3.188)	(-3.133)
(adj) R^2	0.013	0.069	0.178	0.263	0.305
1952 to 1994					
Labor/Disposable	-0.602	-2.470	-5.220	-7.138	-8.430
t-stat.	(-1.386)	(-1.941)	(-2.596)	(-2.918)	(-2.872)
(adj) R^2	0.007	0.050	0.149	0.241	0.303

The table shows the result of the predictive regression $r_{t,t+K} = \alpha + \beta(k)\,\mathbf{x}_t + \varepsilon_{t+K}$ where $\mathbf{x}_t$ is labor-income over disposable income; K is the numbers of quarters ahead and $r_{t,t+K}$ is the cumulative log excess return over K quarters. Number in parenthesis show Newey-West adjusted t-statistics.

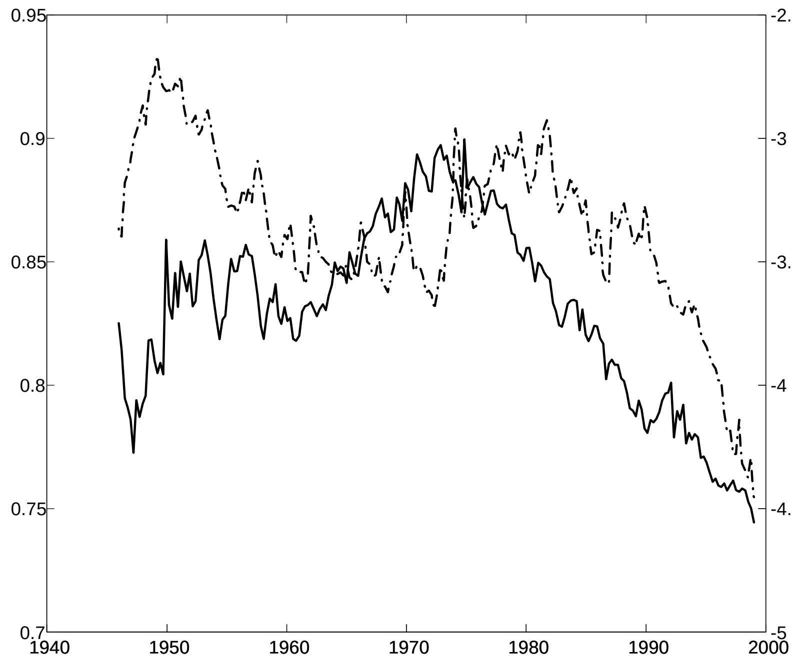
TABLE X

Simulated Distribution of Predictive Regression Coefficients

		Percentiles								
K	x_t	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
1	s_w	-0.486	-0.397	-0.318	-0.241	-0.007	0.225	0.302	0.371	0.473
4	”	-1.879	-1.545	-1.237	-0.943	-0.025	0.868	1.172	1.429	1.807
8	”	-3.659	-2.942	-2.407	-1.84	-0.048	1.687	2.257	2.76	3.525
12	”	-5.285	-4.284	-3.482	-2.684	-0.071	2.52	3.348	4.052	5.024
16	”	-6.73	-5.502	-4.514	-3.507	-0.112	3.323	4.394	5.317	6.548
1	$\ln\left(\frac{D}{P}\right)$	-0.01	-0.007	-0.005	-0.003	0.007	0.028	0.039	0.049	0.063
4	”	-0.039	-0.028	-0.021	-0.011	0.03	0.11	0.151	0.192	0.239
8	”	-0.081	-0.061	-0.041	-0.022	0.059	0.217	0.292	0.37	0.458
12	”	-0.13	-0.094	-0.065	-0.034	0.09	0.318	0.425	0.526	0.653
16	”	-0.183	-0.133	-0.09	-0.048	0.12	0.417	0.551	0.676	0.82
1	$\ln\left(\frac{D}{P}\right)$	-0.008	-0.005	-0.002	0.001	0.02	0.055	0.07	0.084	0.1
	s_w	-0.631	-0.507	-0.416	-0.323	-0.029	0.269	0.365	0.461	0.567
4	”	-0.034	-0.02	-0.009	0.006	0.078	0.211	0.261	0.309	0.374
	”	-2.339	-1.916	-1.583	-1.255	-0.107	1.034	1.395	1.739	2.182
8	”	-0.067	-0.041	-0.017	0.011	0.152	0.396	0.478	0.564	0.662
	”	-4.332	-3.54	-2.986	-2.365	-0.213	1.966	2.625	3.289	3.922
12	”	-0.104	-0.063	-0.027	0.015	0.225	0.555	0.664	0.759	0.876
	”	-6.097	-5.069	-4.282	-3.339	-0.305	2.786	3.727	4.495	5.521
16	”	-0.147	-0.084	-0.038	0.018	0.292	0.697	0.81	0.927	1.051
	”	-7.521	-6.405	-5.315	-4.208	-0.38	3.501	4.691	5.657	6.801
1	$\ln\left(\frac{D}{P}\right) \times s_w$	-0.024	-0.019	-0.015	-0.009	0.017	0.059	0.076	0.09	0.109
4	”	-0.097	-0.075	-0.056	-0.034	0.069	0.228	0.282	0.333	0.391
8	”	-0.196	-0.148	-0.109	-0.066	0.138	0.424	0.518	0.605	0.691
12	”	-0.283	-0.217	-0.162	-0.094	0.203	0.595	0.72	0.83	0.948
16	”	-0.37	-0.286	-0.209	-0.116	0.267	0.752	0.895	1.023	1.153

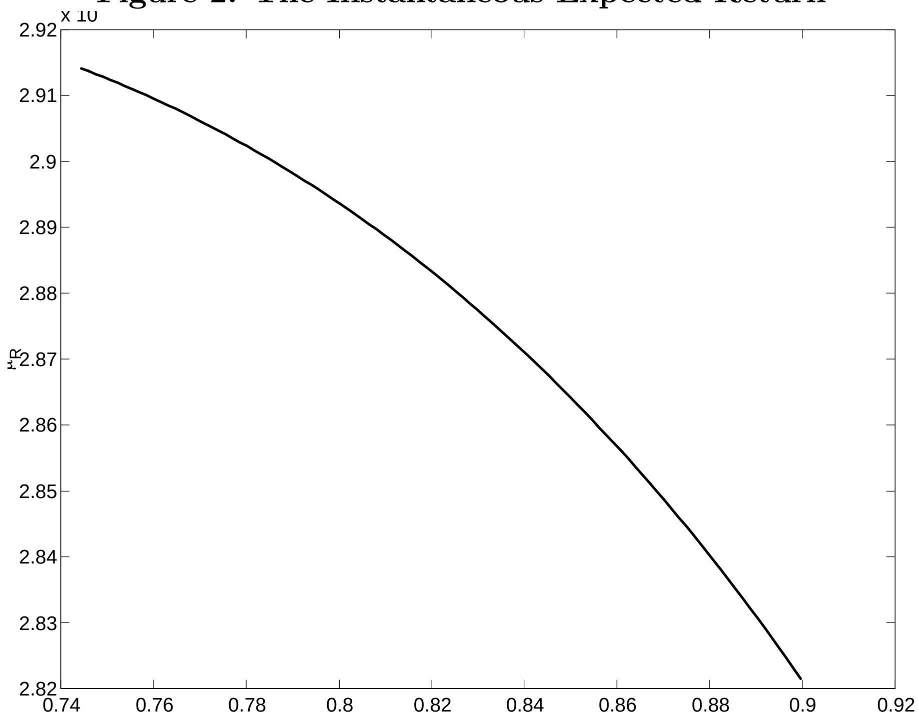
The table shows the distribution of predictive regression coefficients obtained from 10,000 simulations of the system $y_t = a + u_t$, $x_t = \alpha + \phi x_{t-1} + v_t$ where for each simulation, a K period ahead OLS regression is performed. That is, $y_t(K) = \sum_{i=1}^{K} y_{t+i}$ is regressed on x_{t-1}. The parameters a, α and ϕ as well as $\Sigma = E(\varepsilon_t, \varepsilon_t')$ with $\varepsilon_t = (u_t, v_t)'$ are given by their real sample estimates for each regressor x_t as in the table.

Figure 1: Log Dividend Price Ratio and Labor Income to Consumption Ratio



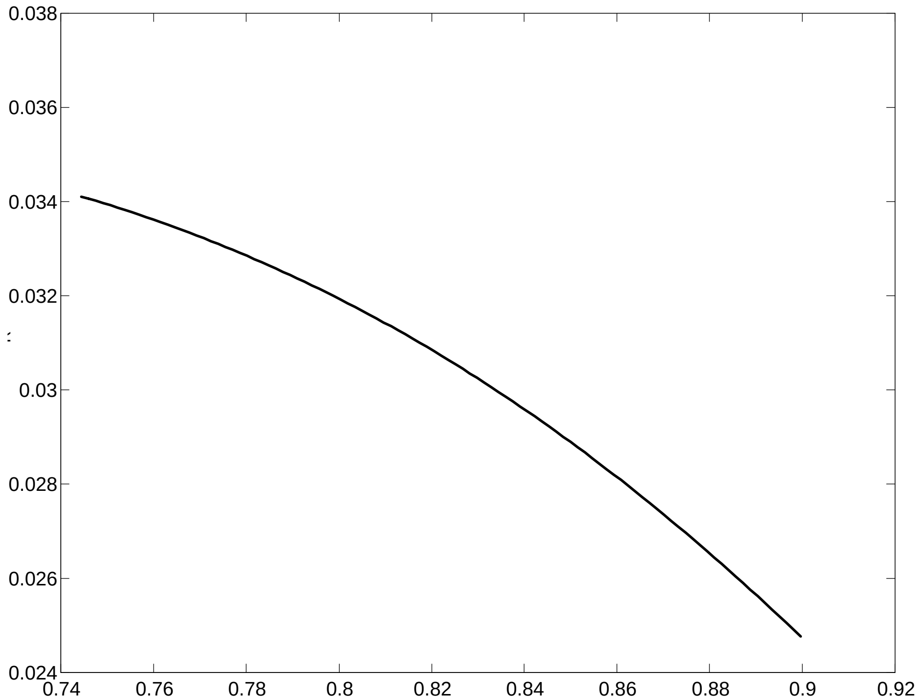
The time series of the labor income - to - consumption ratio (solid line) and of the log dividend price ratio (dash-dotted line).

Figure 2: The Instantaneous Expected Return



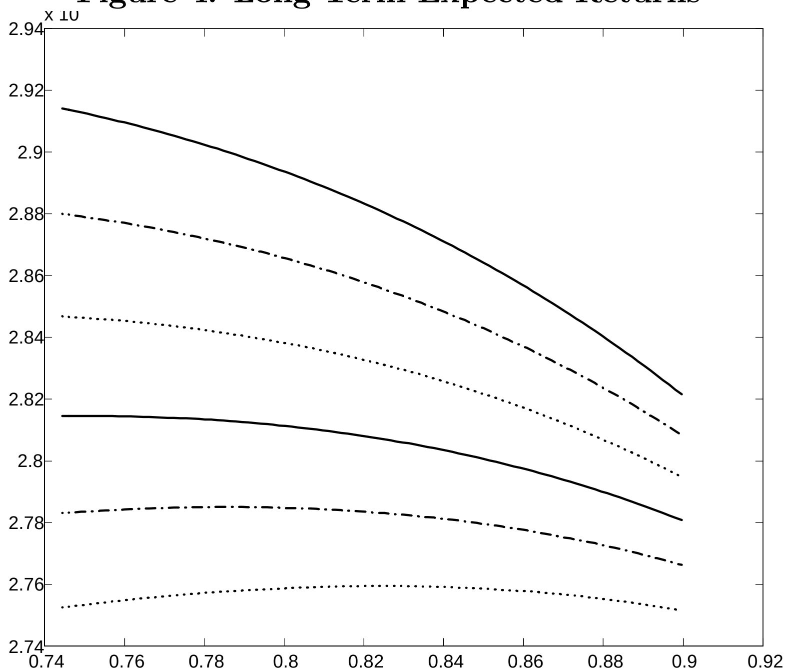
The instantaneous expected excess stock return plotted against the wages-to-consumption ratio $s_w = w/C$.

Figure 3: The Volatility of Returns



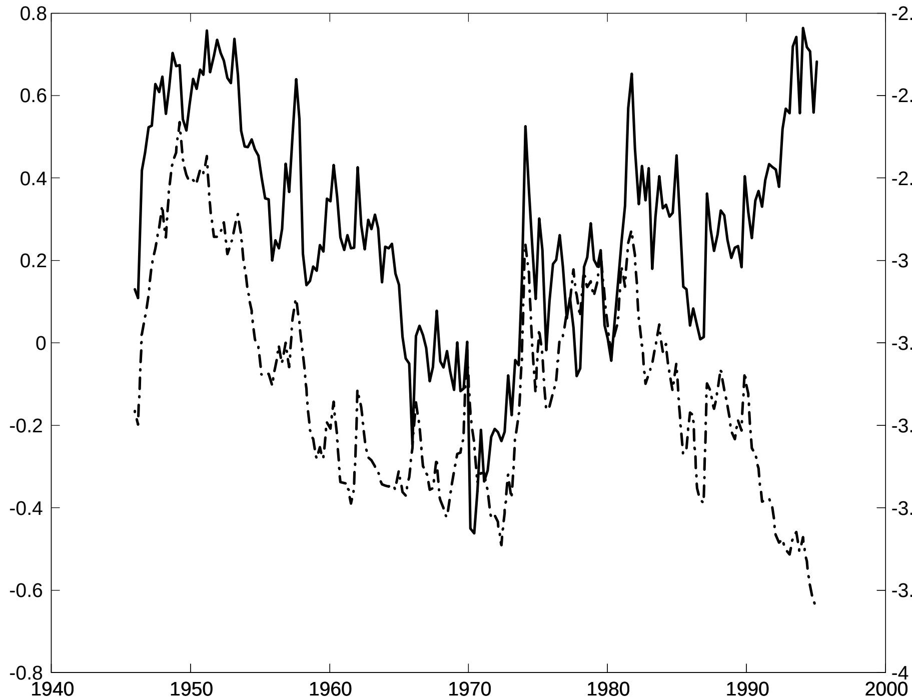
The instantaneous volatility of stock returns plotted against the wages-to-consumption ratio $s_w = w/C$.

Figure 4: Long-Term Expected Returns



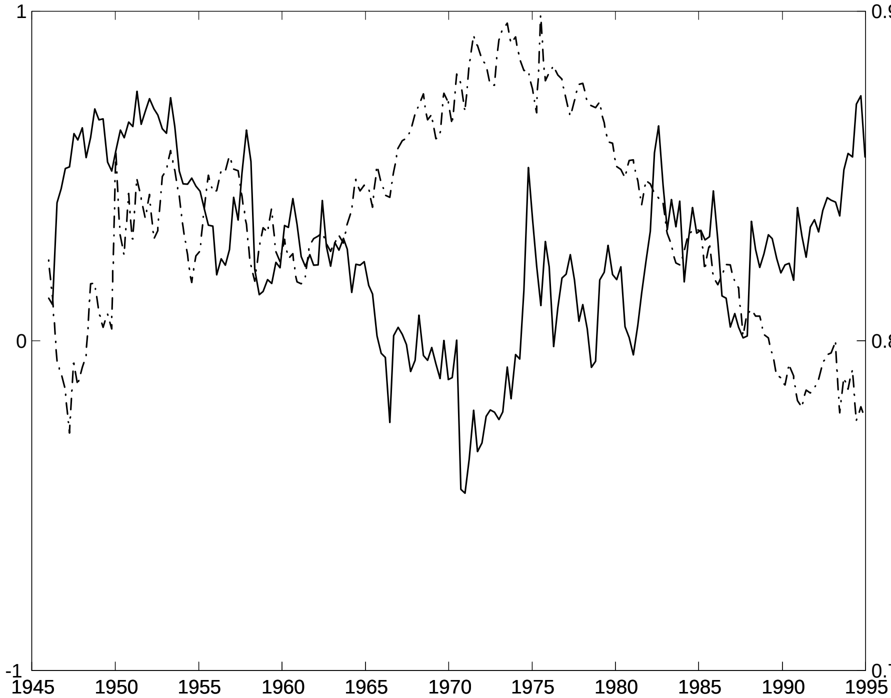
Long term expected returns against wages over consumption s_w for instantaneous, 1, 2, 3, 4 year horizons. Top lines are for shortest horizon.

Figure 5: Long-term Return and Log Dividend Price Ratio



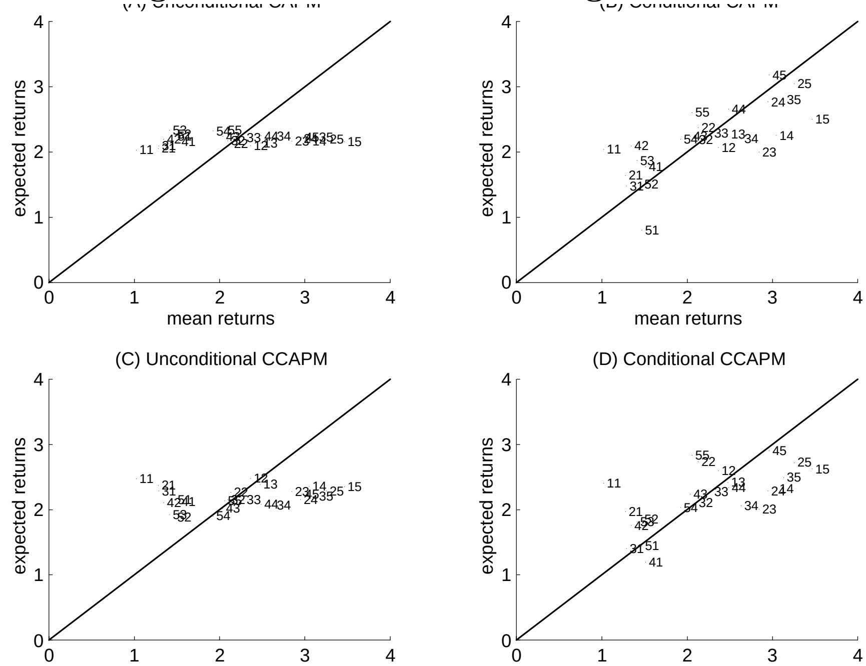
This figure plots the 4 year cumulative return (solid line) lagged by 4 years and the current log dividend price ratio (dash-dotted line)

Figure 6: Long-term Return and Labor Income to Consumption Ratio



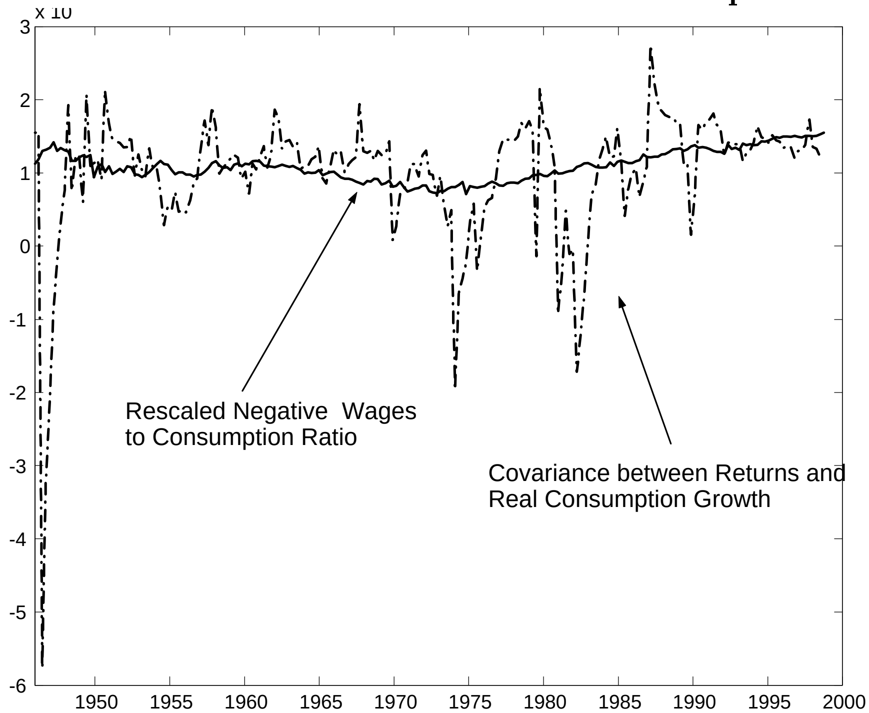
This figure plots the 4 year cumulative return (solid line) lagged by 4 years and the current labor income to consumption ratio (dash-dotted line)

Figure 7: Fitted versus Average Returns



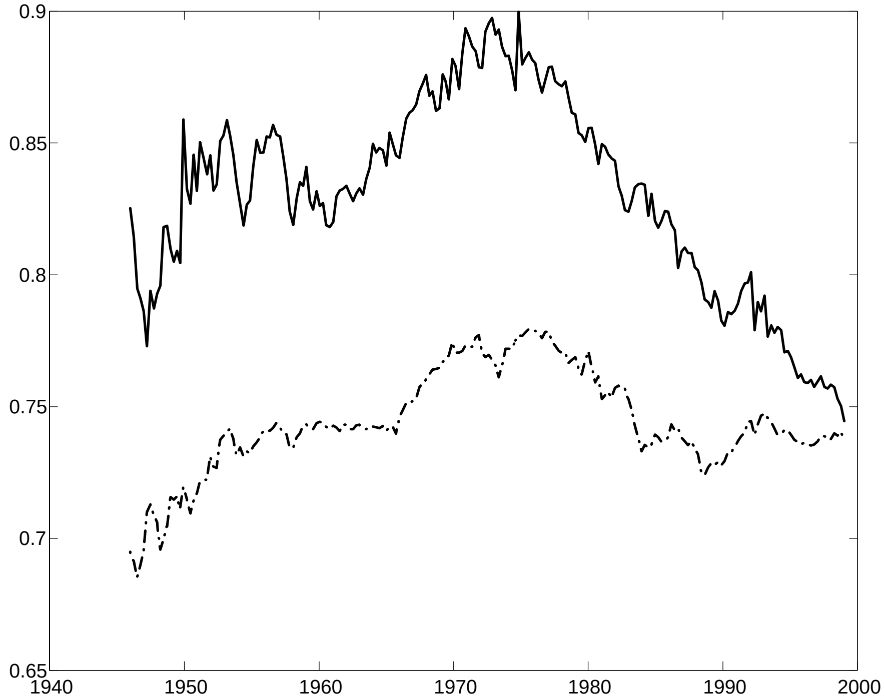
These four plots show the realized versus fitted returns on the 25 FF portfolios. Panel A shows the unconditional CAPM. Panel B shows the conditional CAPM, conditioning by both the share, s_w, and the orthogonalized interaction factor, $R_w \times s_w$. Panel C shows the unconditional CCAPM. Panel D shows the conditional CCAPM, conditioning by both the share s_w, and the orthogonalized interaction factor, $\Delta c \times s_w$.

Figure 8: Covariance between Returns and Consumption Growth



This figure plots the time series of the covariance between returns and real consumption growth (dash-dotted line) along with the negative of the wages to consumption ratio, rescaled to fit the figure.

Figure 9: Time Series of Macro Economic Ratios



Time series of Labor-income to Consumption ratio (solid line) and Labor income to Disposable income ratio (dash-dotted line).