NBER WORKING PAPER SERIES

PRICE DIVIDEND RATIO FACTORS :
PROXIES FOR LONG RUN RISK

Ravi Jagannathan
Srikant Marakani

Working Paper 17484
http://www.nber.org/papers/w17484

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2011

Price Dividend Ratio Factors : Proxies for Long Run Risk
Ravi Jagannathan and Srikant Marakani
NBER Working Paper No. 17484
October 2011, Revised August 2013
JEL No. G11,G12

ABSTRACT

We evaluate the empirical support for a broad class of long run risk models using information in factors extracted through principal component analysis of the covariance matrix of log price dividend ratios of twenty five equity portfolios formed on Size and Book-to-Market. We identify two price-dividend ratio factor proxies for economy wide long run risk, one tracking the volatility of the growth rate in economy wide aggregate consumption, and the other predicting the growth rates in the stock index portfolio dividends and aggregate consumption, consistent with the implications of these models. We show that that the long run risk factor driving expected consumption growth is not recoverable from the cross section of excess returns alone. The price dividend ratio factors perform better than the stock index price dividend ratio and the corporate yield spread, and has information in addition to what is in the slope of the term structure of interest rates, in forecasting the growth rate in real time consumption and stock index dividends. The covariance of excess returns with factor innovations explain the cross section of excess returns on size, book/market, earnings/price ratio, long term reversal, and short term reversal sorted portfolios in a manner robust to look-ahead and useless factor biases. Our findings suggest that the widely used Fama and French (1993) three factor model and the long run risk models studied in the literature are not necessarily inconsistent with each other. They may be representing the same underlying phenomenon, but emphasizing different aspects of reality.

Ravi Jagannathan
Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Leverone/Anderson Complex
Evanston, IL 60208-2001
and NBER
rjaganna@northwestern.edu

Srikant Marakani
Department of Economics and Finance
College of Business
City University of Hong Kong,
P7422, Academic 1
Tat Chee Avenue, Kowloon, Hong Kong
smarakan@cityu.edu.hk

1 Introduction

Recent research in asset pricing has focused on models of dynamic economies in order to provide a better understanding of the underlying economic forces behind various empirical regularities observed in financial markets. Three of the empirical regularities that have received particular attention in the literature are the high historical average risk premium on corporate equities (Mehra and Prescott 1985), large cross sectional variation in historical average returns across various asset classes (Fama and French 1992) (Fama and French 1995) and low risk free rate (Weil 1989).

Among the various representative agent consumption based models that have been proposed for explaining these patterns, three have received wide attention: the external habit formation model of Campbell and Cochrane (1999), the rare disasters model of Rietz (1988) and Barro (2006) and the long run risk model of Bansal and Yaron (2004).[1] Lettau and Wachter (2007) and Santos and Veronesi (2009) argue that it is difficult to reconcile the external habit formation model of Campbell and Cochrane (1999) with the value premium. Gabaix (2008) and Gourio (2008) point out that there is little evidence that value and growth stocks have differential exposure to disaster risk and therefore rare disaster models will have difficulties in explaining the value premium puzzle.

The evidence for the long run risk model has been mixed. On the one hand, Bansal, Yaron, and Kiku (2007), Bansal, Dittmar, and Lundblad (2005), Bansal, Dittmar, and Kiku (2009), Malloy, Moskowitz, and Vissing-Jørgensen (2009) and Da (2009) find that long run risk models can explain the cross section of stock returns. On the other hand, the studies of Beeler and Campbell (2012), Marakani (2009), Ferson, Nallareddy, and Xie (2012), Constantinides and Ghosh (2008) and others either dispute the interpretation of the results of the former studies

---

[1]Asset pricing models can also be written on the basis of production and productivity shocks as in (Zhang 2005). Since our focus is on consumption based representations, we do not examine production based models.

and/or find other empirical evidence casting doubt on the ability of long run risk models to explain the cross section of stock returns.

The mixed empirical support may be in part due to the use of per capita aggregate consumption as a proxy for the consumption of the marginal investor in empirical studies of the long run risk models. There is some evidence suggesting that the consumption of shareholders may be a better proxy. For example, Malloy, Moskowitz, and Vissing-Jørgensen (2009) find that shareholders' consumption risk explains the cross section of stock returns better than the risk associated with aggregate per capita consumption. The difficulty is that shareholder consumption data is available for a shorter length of time and may exhibit non-stationarities as documented by Parker and Vissing-Jorgensen (2010).[2] He and Krishnamurthy (2008) and Adrian, Etula, and Muir (2011) take the stand that financial intermediaries are more likely to consistently be at the margin in securities markets and their consumption and wealth changes may be more relevant for explaining variations in security prices over time and in the cross section.

We therefore assume that the marginal investors' equilibrium consumption evolves according to a long run risk process of a form similar to that proposed by Bansal and Yaron (2004), without imposing the restriction that aggregate per capita consumption is a good proxy for the marginal investors' consumption. We show that the long run risk factors in this model economy are also the factors affecting temporal evolution of the log price to dividend ratios of stocks. Using this fact, the long run risk factors can be estimated through principal component analysis of the covariance matrix of log price to dividend ratios of a collection of stocks. We assume that it is reasonable to ignore the term involving the marginal investors' contempora-

---

[2]Parker and Vissing-Jorgensen (2010) find that the relation between the consumption of high income households, which is a good proxy for shareholder consumption, and aggregate consumption and income fluctuations is non-stationary and exhibits large changes after 1980. Malloy, Moskowitz, and Vissing-Jørgensen (2009) overcome the short data series by extrapolating the shareholder consumption data by using it's projection on other variables but this procedure assumes that the relationships are stationary.

neous consumption growth in the asset pricing relation that we derive, since the state variables and their innovations explain the bulk of the risk premium in the long run risk model.[3]

We use price dividend ratio factors rather than return factors in our analysis as the use of the former enables us to directly estimate the slow moving long run risk state variables. Therefore, their use allows us to test the model implied relationship between these factors and the macroeconomic variables of interest. This would not be possible with only study of the return factors as we show that they do not contain information about the expected consumption/dividend growth factors in the long run risk model. We verify this empirically by showing that this factor, estimated using our methodology, drives out the Fama-French factors in predictive regressions of these quantities.

We also find and clarify the relationship between the asset pricing relations implied by both these methodologies. This relationship implies that, under certain assumptions, both asset pricing relations are approximately equivalent. This finding provides an explanation for the good performance of the Fama-French three factor model as it implies that the latter is approximately equivalent to the price dividend ratio factor model for asset pricing purposes.

Our methodology does not focus on a particular parametric version of the long run risk model, and does not require projecting the future consumption growth and its volatility on the log market price to dividend ratio and the real risk free rate to estimate the long run risk factors as in (Bansal, Yaron, and Kiku 2007), (Constantinides and Ghosh 2008) and (Ferson, Nallareddy, and Xie 2012). This is an advantage since both consumption growth and the real risk free rate are measured with considerable error. As we show using Monte Carlo simulations, this error could be large enough to result in the rejection of the long run risk model too often even when it holds. The disadvantage is that our method will have a lower

---

[3]While the marginal investors' consumption might not be directly observable, it is likely to be positively correlated with measures of aggregate consumption growth. We find that the inclusion of aggregate consumption growth in the empirical analysis does not significantly change our results.

power if the particular parametric model that Constantinides and Ghosh (2008) consider is indeed the right one.

We find that the cross section of average excess returns on several well diversified portfolios are consistent with long run risk models where the long run risk state variables are adequately represented by the two log price dividend ratio factors that we identify using this approach. While this result alone is consistent with a variety of asset pricing models, we also find that the two factors predict future aggregate consumption and dividend growth as well as track contemporaneous consumption growth volatility. This is consistent with long run risk models under the assumption that the factors that affect the consumption of the marginal investor also affects aggregate consumption. Our findings provide some evidence against standard time separable power utility based representative agent models. However, we can not rule out representative agent models with alternative preferences or models where agents have time separable power utility but face decision making costs as in (Jagannathan and Wang 2007) and (Jagannathan, Marakani, Takehara, and Wang 2012). These conclusions are robust to the recent critique of factor models by Kleibergen (2010) and the look ahead bias critique of Ferson, Nallareddy, and Xie (2012).

This enables an answer to one of the important critiques of the long run risk model raised by Beeler and Campbell (2012). Beeler and Campbell (2012) point out that the log market price dividend ratio does not predict dividend or consumption growth. We show that a price dividend ratio factor does predict both dividend growth a real time measure of consumption growth. This does not show up in the log market price dividend ratio as the market price dividend ratio weights most strongly on the volatility price dividend ratio factor.

The rest of the paper is organized as follows. Section 2 introduces our version of the long run risk model which encompasses the ones of Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Zhou and Zhu (2009). Section 3 discusses the implication of the model for the factor structure of log price to dividend ratios and asset pricing. Section 4 describes the

data. Section 5 develops the econometric specifications and discusses the empirical findings. Section 6 concludes.

2　The Long Run Risk Model

Our main aim is to use a model which is flexible enough to encompass a large range of long run risk models extant in the literature while retaining and emphasizing the core assumptions of the framework which are that there are common persistent components of expected consumption and dividend growth and persistent components of volatility. We model the dividends of individual well diversified portfolios as in (Hansen, Heaton, and Li 2008) and (Kiku 2006). The relative advantages of this approach are discussed in detail by Bansal, Dittmar, and Kiku (2009).

We consider the following long run risk model which accommodates the specifications proposed by Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Zhou and Zhu (2009) as special cases.[4] Let c, $X_i, 1 \leq i \leq n$ and $V_j, 1 \leq j \leq m$ be the log consumption process, n processes that determine it's conditional growth rate and m processes that determine the volatility of it's conditional growth rate respectively. Let $d_l, l \leq 1 \leq L$ be the log dividend processes of L portfolios (in general, the lower case variables correspond to the logarithm of the upper case variables). We assume that these quantities follow the processes

$$
\begin{aligned}
c_{t+\Delta t} =& c_t + \left(\mu + \sum_{i=1}^{n} X_{i,t} \right) \Delta t + \sqrt{\sum_{j=1}^{m} \delta_{c,j}^2 V_{j,t}} \left(W_{t+\Delta t} - W_t \right) \\
& - \sum_{k=1}^{m} \varphi_{w,k} \sqrt{V_{k,t}} \left(Z_{k,t+\Delta t} - Z_{k,t} \right)
\end{aligned}
\tag{1}
$$

---

[4]Note that the volatility process has to be modified to an Ornstein-Uhlenbeck one to accommodate the first two specifications. This modification does not affect any of the fundamental theoretical results or empirical analysis.

$$X_{i,t+\Delta t} = X_{i,t}(1 - \alpha_i \Delta t) + \varphi_{i,x} \sqrt{\sum_{j=1}^{m} \delta_{x,i,j}^2 V_{j,t}} \left(Y_{i,t+\Delta t} - Y_{i,t}\right), 1 \le i \le n \qquad (2)$$

$$V_{i,t+\Delta t} = V_{i,t} - \kappa_i(V_{i,t} - \bar{V}_i)\Delta t + \sigma_i \sqrt{V_{i,t}} \left(Z_{i,t+\Delta t} - Z_{i,t}\right), 1 \le i \le m \qquad (3)$$

$$
\begin{aligned}
d_{l,t+\Delta t} =& d_{l,t} + \left(\mu_l + \sum_{i=1}^{n} \phi_{l,i} X_{i,t}\right)\Delta t + \pi_{l,c}\left(\Delta c_{t+\Delta t} - \left(\mu + \sum_{i=1}^{n} X_{i,t}\right)\Delta t\right) \\
&+ \sum_{i=1}^{n} \pi_{i,l,x}(X_{i,t+\Delta t} - X_{i,t}(1 - \alpha_i \Delta t)) \\
&+ \sum_{j=1}^{m} \pi_{j,l,w}\sigma_j \sqrt{V_{j,t}} \left(Z_{j,t+\Delta t} - Z_{j,t}\right) \\
&+ \sqrt{\sum_{k=1}^{m} \delta_{l,d,k}^2 V_{k,t}} \, \sigma_{l,d}(B_{l,t+\Delta t} - B_{l,t})
\end{aligned}
\qquad (4)
$$

where W, Y_i, $1 \le i \le n$, Z_j, $1 \le j \le m$ and B are independent Wiener processes and $\sum_{i=1}^{m} \delta_{c,i}^2 = \sum_{j=1}^{m} \delta_{x,i,j}^2 = \sum_{k=1}^{m} \delta_{l,d,k}^2 = 1$ (as pointed out by Zhou and Zhu (2009), these variables are necessary to ensure that the market volatility is decoupled from the consumption growth volatility as is the case in the data). The basic time interval of the process is assumed to be the same as that for which consumption is observed. This ensures that the stochastic discount factor can be related to the innovations in the processes c, X and V. If the basic time interval of the process is smaller than that for which consumption is observed, time aggregation effects prevent the calculation of the stochastic discount factor as shown by Bansal, Yaron, and Kiku (2007) as the true innovations cannot be recovered from the observed data. We note that at least some of the α_i and ν_i must be small for the risks to be long lived and carry a high price.

In the above equations, consumption is defined as a rate rather than per period so that consumption from time t to $t + \Delta t$ is $C_{t+\Delta t}\Delta t$ and log consumption from t to $t + \Delta t$ is $c_{t+\Delta t} + \log \Delta t$. While this generally makes no difference as the log consumption is just offset by a constant, it ensures that the continuous time limit exists and (as we show in appendix A) also makes it easy to obtain it. It further shows that the solution method of Bansal and

Yaron (2004) is general enough to apply to continuous time models. One consequence of this definition is that the log-linearization constants depend on Δt. This is due to the fact that consumption and dividends are flow variables whose magnitude depend on the time interval (the shorter the time interval, the smaller the consumption and dividend). This fact implies that the log-linearization constants, which are functions of the average wealth to consumption or price to dividend ratio, are inversely related to the time scale. This explains the dependence of these log-linearization constants, and hence the market prices of risk, on the time unit chosen in the formulae below.[5]

This long run risk process, when written in continuous time, incorporates the one proposed by Zhou and Zhu (2009) as a special case (specifically with $n = 1$, $m = 2$ and $\varphi_{w,i} = 0, 1 \leq i \leq m$). When the volatility process (3) is modified to an Ornstein-Uhlenbeck one plus a constant by modifying the second term to $\sigma_{i,w}\Delta Z_{i,t+1}$ and the final term in (1) to $-\sum_{k=1}^{m} \varphi_{w,k}(Z_{k,t+\Delta t} - Z_{k,t})$, it incorporates the ones proposed by Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007) as special cases (specifically with $n = 1$, $m = 1$, $\varphi_{w,i} = 0, 1 \leq i \leq m$, $\pi_{i,l,x} = 0, 1 \leq i \leq n, 1 \leq l \leq L$ and $\pi_{i,l,w} = 0, 1 \leq i \leq m, 1 \leq l \leq L$).

We now look more closely at the dividend growth specification (4) in this formulation. Most importantly, we note that the embedded assumption in (4) that the expected dividend growth of all the portfolios depends on the same set of state variables X that determine expected consumption growth is not essential but convenient. In general, we can make the expected dividend growth depend on more AR(1) state variables Y so that

$$E_t[\Delta d_{l,t+\Delta t}] = \left(\mu_l + \sum_{i=1}^{n} \phi_{l,i} X_{i,t} + \sum_{j=1}^{n_o} \zeta_{l,j} Y_{j,t}\right) \Delta t \tag{5}$$

instead of $(\mu_l + \sum_{i=1}^{n} \phi_{l,i} X_{i,t})\Delta t$. The innovations of these state variables Y which are unrelated to future consumption/dividend growth or consumption growth volatility must then have zero

---

[5]This makes a small contribution as the solution method of Bansal and Yaron (2004) is simpler than the continuous time solution method used by Zhou and Zhu (2009)

price of risk and this can be empirically examined. A more intuitive way to think about this is to interpret the restriction of using a single set of state variables to determine both expected quantities as a criterion for the choice of portfolios to use. In other words, we want to choose portfolios whose characteristics are related to future consumption growth. To do so effectively, we use the studies of Parker and Julliard (2005) and Malloy, Moskowitz, and Vissing-Jørgensen (2009) to help pin-point assets whose cross-sectional variation is related to future consumption growth. Once the choice is made, the extra restrictions on any identified state variables corresponding to Y can then be examined.[6]

Menzly, Santos, and Veronesi (2004) develop an alternative way of modeling the process for aggregate dividends starting from the process for the dividends of individual firms in an economy. While attractive for theoretical analysis, the disadvantage in empirical studies is that it ignores investment which, as shown by Bansal, Dittmar, and Kiku (2009), plays a major part in linking the aggregate stock market dividends to the dividends paid out by individual firms.

Consumers in the model have Epstein-Zin-Weil preferences (Epstein and Zin 1989) (Weil 1990)

$$U_t = ((1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta E_t[U_{t+1}^{1-\gamma}]^{\frac{1}{\theta}})^{\frac{\theta}{1-\gamma}} \tag{6}$$

As noted by Bansal and Yaron (2004), we need $\gamma > 1/\psi$ to generate a positive equity risk premium as expected dividend growth is positively related to expected consumption growth (as noted by Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007), Bansal, Dittmar, and Lundblad (2005), Bansal, Dittmar, and Kiku (2009) and others). This implies that they prefer early resolution of uncertainty and that shocks to expected consumption growth carry a positive price of risk (as pointed out by Kaltenbrunner and Lochstoer (2010)) which is high if the expected consumption growth is persistent. This high price of risk results in a high equity premium and low risk-free rate.

---

[6] With our choice of portfolios, we do not find any state variables or factors of this type.

3 Factor structure of log P/D ratios

In appendix A, we show, using the approach of Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007), that this long run risk model implies that

$$\log\left(\frac{P_{l,t}}{D_{l,t}}\right) = p_{l,t} - d_{l,t} = A_{0,l} + \sum_{i=1}^{n} A_{1,l,i} X_{i,t} + \sum_{j=1}^{m} A_{2,l,j} V_{j,t} \tag{7}$$

where $P_{l,t}$ is the price of portfolio l, $A_{1,l,i} = \frac{(\phi_{l,i} - 1/\psi)\Delta t}{1 - \nu_{1,l}(1 - \alpha_i \Delta t)}, 1 \leq i \leq n$ ($\nu_{1,l}$ being a log-linearization constant which is endogenously determined in the model) and where the expressions for $A_{2,l,j}, 1 \leq j \leq m$ are derived in appendix A. This generalizes the equivalent results by Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Zhou and Zhu (2009) to the situation when there are multiple state variables describing predictable consumption growth and consumption growth volatility. If additional state variables Y are needed in determining the dividend growth, (7) changes to

$$\log\left(\frac{P_{l,t}}{D_{l,t}}\right) = p_{l,t} - d_{l,t} = A_{0,l} + \sum_{i=1}^{n} A_{1,l,i} X_{i,t} + \sum_{k=1}^{n_0} A_{1o,k,l} Y_{k,t} + \sum_{j=1}^{m} A_{2,l,j} V_{j,t} \tag{8}$$

where $A_{1o,l,k} = \frac{\zeta_{l,k}\Delta t}{1 - \nu_{1,l}(1 - \alpha_k \Delta t)}, 1 \leq k \leq n_0$ where α_k is the AR(1) coefficient for the process for $Y_{k,t}$.

Since the real risk-free rate can be viewed as a special type of dividend-price ratio, it also follows that

$$r_{f,t} = A_{0,f} + \sum_{i=1}^{n} A_{1,f,i} X_{i,t} + \sum_{j=1}^{m} A_{2,f,j} V_{j,t} \tag{9}$$

where $A_{1,f,i} = 1/\psi$.

We show that the log stochastic discount factor for this model is given by

$$
\begin{aligned}
m_{t+\Delta t} =\Delta t & \left(\Gamma_0 + \sum_{i=1}^{n} \Gamma_{1,i} X_{i,t} + \sum_{j=1}^{m} \Gamma_{2,j} V_{j,t} \right) \\
& - \alpha_c \sqrt{\sum_{j=1}^{m} \delta_{c,j}^2 V_{j,t}} (W_{t+\Delta t} - W_t) \\
& - \sum_{i=1}^{n} \alpha_{x,i} \sqrt{\sum_{j=1}^{m} \delta_{x,i,j}^2 V_{j,t}} (Y_{i,t+\Delta t} - Y_{i,t}) \\
& - \sum_{j=1}^{m} \alpha_{v,j} \sqrt{V_{j,t}} (Z_{j,t+\Delta t} - Z_{j,t})
\end{aligned}
\tag{10}
$$

where $\Gamma_{1,i} = -1/\psi$, $\alpha_c = \gamma$ and $\alpha_{x,i} = \frac{\gamma - 1/\psi}{1 - \nu_1 (1 - \alpha_i \Delta t)}$.[7] The expression for $\alpha_{v,j}$ is complicated and does not directly concern us here, but we note that it was shown by Bansal and Yaron (2004) that $\alpha_{v,j} < 0$ if $\gamma - 1/\psi > 0$ and $\psi > 1$.

The relatively simple form of $\lambda_{x,i}$ implies that it can be used together with a reasonable approximation for $1 - \nu_1 = \frac{\exp(\overline{c-w})\Delta t}{1 + \exp(\overline{c-w})\Delta t} \approx \exp(\overline{c-w})\Delta t$ to estimate $\gamma - 1/\psi$ once a component X_i is identified. The estimation of ν_1, which can at best be done heuristically, is a cost that has to be paid when the parameters are not explicitly specified. It must be cautioned that while this estimate is likely to be imprecise due to it's indirect nature, it is still useful in that it allows to relate the empirical results back to the underlying preferences.

In the case of no measurement error, (7) can be inverted to express the state variables (X_i, V_j) as a linear combinations of log P/D ratios. This enables the expression of the log stochastic discount factor as

$$
m_{t+\Delta t} = \Delta t \left(\tilde{\Xi}_0 + \sum_{i=1}^{m+n} \tilde{\Xi}_{1,i} F_{i,t} \right) - \alpha_c (c_{t+\Delta t} - c_t) - \sum_{i=1}^{m+n} \alpha_{q,i} IF_{i,t+\Delta t}
\tag{11}
$$

---

[7]Note that the value of $\alpha_{x,i}$ depends on Δt but the risk premium does not. $\alpha_{x,i}$ varies inversely with Δt as $1 - \nu_1$ is proportional to Δt. Since the risk premium due to this risk is given by the product of $\alpha_{x,i}$ and the covariance between the return and the innovation to X_i which is proportional to Δt, the inverse relationship between $\alpha_{x,i}$ and Δt implies that the risk premium is independent of it.

where F_i and IF_i, $1 \leq i \leq n+m$ are the $n+m$ principal components of the log P/D ratios (or, equivalently, any linear combination of $n+m$ log P/D ratios) and their innovations respectively.

(7) implies that the log P/D ratios of assets follow a strict factor structure (up to the loglinear approximation) in the model.[8] Since log P/D ratios are not exact linear combinations of a small number of factors in the data, we use a slightly modified relation in our empirical work. This relation is

$$\log\left(\frac{P_{l,t}}{D_{l,t}}\right) = p_{l,t} - d_{l,t} = A_{0,l} + \sum_{i=1}^{n} A_{1,l,i}X_{i,t} + \sum_{j=1}^{m} A_{2,l,j}V_{j,t} + \epsilon_{l,t} \tag{12}$$

where $\epsilon_{l,t} \sim N(0, V_e)$ are i.i.d. $\epsilon_{l,t}$ can be thought of as deviations that arise due to market imperfections such as illiquidity or due to the existence of incompletely diversified idiosyncratic factors. In section 5, we show that, given the assumed error structure, principal component analysis (or singular value decomposition) can be used to estimate the linear subspace that the $n+m$ factors span once $n+m$ is specified and that statistical tests suggested in the literature can be used to estimate $n+m$ from the data.

This differs from the methodology used by Bansal, Yaron, and Kiku (2007) and Ferson, Nallareddy, and Xie (2012) in estimating the linear subspace of the factors with the use of several log P/D ratios rather than the projection of the realized long term consumption growth and its volatility on the log market P/D ratio and the real risk free rate. We show in appendix C, by using Monte Carlo simulations of the long run risk model, that our methodology produces much fewer spurious rejections of the model when reasonable measurement errors in consumption growth and the real risk free rate are taken into account.

The principal asset pricing relation that we use in this study is

$$E_t\left[\exp\left(\tilde{\Xi}_0 + \sum_{i=1}^{m+n} \tilde{\Xi}_{1,i}F_{i,t} - \alpha_c \Delta c_{t+1} - \sum_{i=1}^{m+n} \alpha_{q,i}IF_{i,t+1}\right)(R_{i,t+1} - R_{f,t+1})\right] = 0 \tag{13}$$

---

[8]The reader must note that these factors are different though related to the pricing factors discussed below. This terminology for both types of quantities is standard but unfortunate.

derived using the standard asset pricing relation

$$E_t[\exp(m_{t+1})(R_{i,t+1} - R_{f,t})] = 0 \tag{14}$$

and (11) (with Δt set to 1).

We empirically evaluate the asset pricing relation in equation (14) using cross sectional regressions and GMM. The cross sectional regression results are particularly useful in relating our results to those in the literature that use various linear beta pricing representations. The appropriate pricing relations for the cross sectional regressions are developed in the next section. In contrast, the GMM results directly examine the nonlinear pricing relationship (14).

3.1 The Linear Beta Pricing Relation

The beta pricing relation can be derived from the fundamental equations of the model (1), (2), (3) and (4) together with the form of the preferences (6). To simplify the notation, we set Δt to 1 in this derivation and use the index l for the asset rather than i for clarity as the latter is generally used as the index for the X state variables.

Theorem 1. *The model defined by (37), (38), (39) and (40) implies the beta pricing relation*

$$\begin{aligned}
E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t} - \sum_{j=1}^{m} \beta_{l,V_j} V_{j,t}\right] &= \\
\sum_{j=1}^{m}\left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^2\varphi_{x,i}^2 + \alpha_{v,j}\beta_{l,Z_j}\sigma_j^2 + \alpha_c\beta_{l,c}\delta_{c,j}^2\right)\bar{V}_j & \tag{15} \\
= \sum_{i=1}^{n} \beta_{l,Y_i}\lambda_{Y_i} + \sum_{j=1}^{m} \beta_{l,Z_j}\lambda_{Z_j} + \beta_{l,c}\lambda_c &
\end{aligned}$$

13

where

$$\lambda_{Y_i} = \alpha_{x,i}\varphi_{x,i}^2 \sum_{j=1}^{m} \delta_{x,i,j}^2 \bar{V}_j \tag{16}$$

$$\lambda_{Z_j} = \alpha_{v,j}\sigma_j^2 \bar{V}_j \tag{17}$$

$$\lambda_c = \alpha_c \sum_{j=1}^{m} \delta_{c,j}^2 \bar{V}_j \tag{18}$$

where the $\beta_{l,V_j}, 1 \leq j \leq m,\ \beta_{l,Y_i}, 1 \leq i \leq n,\ \beta_{l,Z_j}, 1 \leq j \leq m$ *and* $\beta_{l,c}$ *are, respectively, the multiple regression coefficients obtained on regressing* $r_{l,t+1} - r_{f,t}$ *on* $V_{j,t}, 1 \leq j \leq m,$ $Y_{i,t+1}, 1 \leq 1 \leq n,\ Z_{j,t+1}, 1 \leq j \leq m$ *and the residuals obtained on regressing* Δc_{t+1} *against* $X_{i,t}, 1 \leq i \leq n.$

This can be written in terms of the $n+m$ *log price dividend ratio factors* $F_i, 1 \leq i \leq m+n$ *and their innovations* $IF_i, 1 \leq n+m$ *as*

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\mathrm{Var}\left[r_{l,t+1} - r_{f,t} - \sum_{i=1}^{m+n} \beta_{l,F_i}F_{i,t}\right] = \\ \sum_{i=1}^{m+n} \beta_{l,IF_i}\lambda_{IF_i} + \beta_{l,c}\lambda_c \tag{19}$$

When there is only one volatility factor, the beta pricing relation can also be written as

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\mathrm{Var}\left[r_{l,t+1} - r_{f,t}\right] \\ = \beta_{l,V}^2 \mathrm{Var}[V_t] + \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i}^2\varphi_{x,i}^2 + \alpha_v\beta_{l,Z}\sigma^2 + \alpha_c\beta_{l,c}\delta_c^2\right)\bar{V} \tag{20} \\ = \beta_{l,V}^2 \lambda_V + \sum_{i=1}^{n} \beta_{l,Y_i}\lambda_{Y_i} + \beta_{l,Z}\lambda_Z + \beta_{l,c}\lambda_c$$

with the βs defined in a similar manner as above and where

$$\lambda_V = \mathrm{Var}[V_t] \tag{21}$$

$$\lambda_{Y_i} = \alpha_{x,i}\varphi_{x,i}^2 \bar{V} \tag{22}$$

$$\lambda_Z = \alpha_v \sigma^2 \bar{V} \tag{23}$$

$$\lambda_c = \alpha_c \bar{V} \tag{24}$$

This can be written in terms of the $n+m$ log price dividend ratio factors $F_i, 1 \leq i \leq n+1$ and their innovations $IF_i, 1 \leq n+1$ as

$$
\begin{aligned}
E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\mathrm{Var}\left[r_{l,t+1} - r_{f,t}\right] \\
= \beta_{l,F_k}^2 \lambda_{F_k} + \sum_{i=1}^{n+1} \beta_{l,IF_i} \lambda_{IF_i} + \beta_{l,c}\lambda_c
\end{aligned}
\tag{25}
$$

where the βs are as defined above and F_k is some log price dividend ratio factor that has non-zero correlation with the volatility factor.

Proof. See appendix B. □

In addition to the model considered in this paper, (19) also holds for the long run risk models of Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007), Yang (2011) and Zhou and Zhu (2009). In addition, (25) holds for the long run risk models of Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Yang (2011). (It does not hold for the model of Zhou and Zhu (2009) as it has more than one volatility factor.) Proof of these statements are available upon request.

Theorem 2. *The model defined by (1), (2), (3) and (4) implies the beta pricing relation*

$$E[r_{l,t+1} - r_{f,t}] = \sum_{j=1}^{m} \beta_{l,V_j} \bar{V}_j \tag{26}$$

15

where β_{l,V_j} is the coefficient obtained upon regressing $r_{l,t+1} - r_{f,t}$ on $V_{j,t}$.

This can be written in terms of the $n + m$ log price dividend ratio factors $F_i, 1 \leq i \leq m + n$ as

$$E[r_{l,t+1} - r_{f,t}] = \sum_{i=1}^{m+n} \beta_{l,F_i} \lambda_{F_i} \tag{27}$$

Proof. See appendix B. $\qquad\qquad\square$

(27) holds for our model because of the Bessel process used for the volatility. It does not therefore hold for the long run risk models of Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007) which use Ornstein-Uhlenbeck processes for the volatility. While it can be argued that a Bessel process is a more reasonable one from a theoretical point of view as it does not admit negative values,[9] this means that (27) is not a general implication of long run risk models. However, it's analysis is interesting and can provide some insight as to the economic forces that determine the time series variation in the cross section of expected returns.

It is also interesting to note that, when there is only volatility factor (as we find in this study), (27) is a statement of the fact that there is only one factor driving the cross section of expected returns. In this sense, the results of Gibbons and Ferson (1985) are consistent with our finding of one volatility factor as they find that the returns of the Dow Jones 30 stocks are consistent with a single, time-varying risk premium.

We are unable to test this hypothesis satisfactorily in our approach as the lagged factor betas are not significantly different from each other. This is related to econometric issues which arise due to the fact that the major part of the return is composed of the long run risk factor innovations. This leads to poor estimation of the betas with respect to the lagged factors and spurious rejections of the relation in asset pricing tests. This can be shown using Monte Carlo simulations.[10]

---

[9]In the calibration of Bansal, Yaron, and Kiku (2012), the probability of negative volatilities in the stable distribution is nearly 20%.

[10]The analysis of a Monte Carlo simulation showing this in the context of the CAPM is available upon request.

3.2 Relation to Standard Linear Factor Models

We note from the standard loglinear approximation of Campbell and Shiller (1988) that innovations of the log P/D ratios are similar to innovations in excess returns minus dividend growth.

Explicitly, the loglinear approximation (28) states that the return is approximately given by

$$r_{l,t+1} - r_{f,t} \approx \nu_{0,l} - r_{f,t} + \nu_{1,l}(p_{l,t+1} - d_{l,t+1}) - (p_{l,t} - d_{l,t}) + \Delta d_{l,t+1} \tag{28}$$

This, in turn, can be written as

$$
\begin{aligned}
r_{l,t+1} - r_{f,t} \approx & E_t[r_{l,t+1} - r_{f,t}] + \underbrace{\nu_{1,l}(p_{l,t+1} - d_{l,t+1} - E_t[p_{l,t+1} - d_{l,t+1}])}_{price\ dividend\ ratio\ innovation} \\
& + \underbrace{(d_{l,t+1} - d_{l,t} - E_t[d_{l,t+1} - d_{l,t}])}_{dividend\ growth\ innovation}
\end{aligned}
\tag{29}
$$

Hence, the factor structure of excess returns is composed of three parts

1. the factor structure in expected excess returns (given by the V factors);

2. the factor structure of innovations to the log price dividend ratios;

3. the factor structure in dividend growth innovations; and

Since the model does not constrain the factor structure in dividend growth innovations, it does not imply any specific relation between the return and P/D factor structures. We do, however, show in appendix C that it can be chosen in a manner to make these factor structures consistent with data. We also document, in this appendix, that these factor structures are, in fact, different in the data so that the use of one instead of the other is not vacuous. The dividend growth innovation factor structure, loosely speaking, can thus explain the fact that the price dividend ratios of the twenty five Fama-French portfolios have a two factor structure while their excess returns have a three factor structure.

17

In other words, the methodology that we use closely parallels the standard factor analysis and principal component analysis methods of Connor and Korajczyk (1986), Lehmann and Modest (1988), Lehmann and Modest (2005) and Connor and Korajczyk (2009) but differs in the way the factors are constructed. While standard factor analysis constructs factors from the returns themselves, this methodology pays more attention to returns that are not explained by contemporaneous dividend growth, or in other words, to the more interesting non-trivial part of returns.

It must be emphasized that, despite these close parallels, the methodology used by us is vital as it enables to estimate the slow moving long run risk factors directly which is not possible with the return factors. This is because the return factors only directly contain the volatility type long run risk factors and even these are swamped by the remaining sets of factors identified above, namely the innovations in the long run risk factors and the dividend growth innovations.

4   Data

We examine the cross section of annual returns on several stock portfolios for the period 1943 to 2008, i.e., 66 years. While we primarily use annual consumption data, we compute consumption growth volatility using quarterly consumption data to obtain more precise estimates. Consumption data is obtained from the National Income and Product Accounts (NIPA) tables available at the BEA web site. Real annual per capita consumption is defined to be the nominal aggregate annual consumption of nondurables and services divided by the NIPA estimate of the mid-year population and deflated by the implicit personal consumption deflator.[11] Annual consumption growth is defined to be the first difference of the logarithm of this series.

---

[11]Since we make use of data expressed in terms of chained dollars, we use a Tornqvist type index (Whelan 2000) to construct the implicit consumption deflator.

Quarterly seasonally adjusted consumption data is also obtained from NIPA and it's growth is defined in an analogous manner.

The proxy for the nominal risk-free rate is the Fama 3 month T-bill rate obtained from CRSP. It is converted to two proxies of the real risk free rate using the realized and past inflation as measured by the CPI (as discussed in the relevant section of this paper). The CPI data for this is obtained from CRSP.

The stock market proxy (used to determine the relationship between the factors and future dividend growth and expected returns) is defined as the CRSP value-weighted index of all stocks listed on the NYSE, AMEX and NASDAQ. The construction of portfolios based on size and book-to-market ratios is as in (Fama and French 1993) and (Fama and French 1996). Data on the 25 (5×5) portfolios sorted on the basis of both these characteristics is obtained from Ken French's web site.

For testing the asset pricing relationships with portfolios other than the ones used to estimate the factors (we call this out of sample testing), we use three sets of ten portfolios each formed on the basis of long term reversal, short term reversal and the earnings to price (E/P) ratio. The long term reversal portfolios are formed monthly on the basis of stock's return over the past five years minus it's return over the past year. In other words, they are formed at time $t - 1$ (time being indexed by month) by sorting stocks into ten portfolios according to their returns from $t - 61$ to $t - 13$. Similarly, short term reversal portfolios are formed at time $t - 1$ by sorting stocks into ten portfolios based on their return from $t - 2$ to $t - 1$. The E/P based portfolios are formed at the end of June of year t by sorting stocks into ten portfolios (using NYSE breakpoints) on the basis of their E/P ratios where E is defined to be the earnings before extraordinary items during fiscal year $t - 1$ and P is defined to be the market capitalization at the end of December of year $t - 1$. Data on these thirty portfolios is also obtained from Ken French's web site.

Monthly dividends of these portfolios are calculated using the difference between the returns of the corresponding dividend reinvested and non-reinvested portfolios. The price-dividend ratios are then calculated by dividing the real price of the non-reinvested portfolio by the sum of the lagged twelve real monthly dividends. This procedure accounts for the pronounced seasonality of the dividend series. The nominal prices and dividends are deflated by the CPI to get these real prices and dividends. As pointed by Van Binsbergen and Koijen (2010), the effect of assumptions regarding the handling of dividends paid during the year on the price dividend ratios is negligible with the correlation between the different measures being about 0.9999.

Real time consumption data is obtained from the web site of the Federal Reserve Bank of St. Louis and is described by Croushore and Stark (2001). The real consumption during a quarter is defined to be the sum of the real consumption of nondurables and services during that quarter. The real consumption during a year is defined to be the sum of the real consumptions during each quarter of that year. Real per capita consumption during a period is defined to be the real consumption during that period divided by the mid-period estimate of population. The real time annual per capita consumption growth for year t is defined to be the difference between the logarithms of the real per capita consumptions during years t and $t-1$ respectively as calculated using data of vintage Q1 of year $t + 1$. To provide an example, the data set of Q1 1976 vintage is used to construct the real time annual per capita consumption growth for 1975. It is constructed by adding the real nondurables and services consumptions of Q1-Q4 1974 and Q1-Q4 1975, dividing each of them by the mid-year estimates of the population, and then taking the difference of the logarithms of the corresponding quantities.

5 Empirical Findings

5.1 Structural Break Implied by the Factors

Marakani (2009) documents strong evidence that the parameters of long run risk models could not have been the same before and after 1942. Hence, we only consider the post 1942 period in our analysis and assume that consumers are myopic and do not consider the possibility of regime change in the model.[12] We defer the examination of an extended model where the consumers are aware of possible regime shifts to future research.[13]

5.2 Construction of the Principal Components and Their Innovations

From (7), the problem of obtaining the factors of log P/D ratios is, for a fixed number of factors $n + m$, equivalent to the problem of finding time series processes $F_{i,t}^{n+m}$ to solve

$$V(n+m) = \min_{\Lambda, F^{n+m}} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{i,t} - \Lambda_i^{n+m} F_{i,t}^{n+m} \right)^2 \tag{30}$$

where X is the matrix of demeaned log P/D ratios, N is the number of portfolios, T is the length of the time series, F are the factors and Λ are the loadings of the individual log P/D ratios on them (the superscript $n + m$ keeps track of the number of assumed factors). The equivalency of the two problems follows trivially from the assumption that the error terms are i.i.d and Gaussian. Hence, this problem is the same as the well studied standard factor analysis problem (of which (Connor and Korajczyk 2009) is an excellent review). The assumptions regarding the error terms are not crucial for our results as they hold even if we perform the principal component analysis after first scaling the log P/D ratios to make them each have unit variance or after first scaling them each according to their residual variances. In other words, our results are robust to the use of different specifications for the error term.

---

[12]The use of post-1945 or post-1950 data does not significantly change our results.

[13]In this context, we note that Bekaert and Engstrom (2010) have recently argued that habit formation models are better able to incorporate the very different dynamics observed during and after the Great Depression.

Hence, the factors can be calculated by singular value decomposition of the matrix of demeaned log P/D ratios. This is equivalent to the more usual method of using the eigenvectors of the covariance matrix or directly solving (30), but is preferred because it has greater numerical stability. The number of relevant factors $k = m + n$ is determined by using the information criterion

$$\operatorname*{argmin}_{k} \mathrm{IC}_{p2} \equiv \operatorname*{argmin}_{k} \left(\log V(k) + 2k \left(\frac{N+T}{NT} \right) \log \min(N, T) \right) \tag{31}$$

suggested by Bai and Ng (2008) and by Connor and Korajczyk (2009). This method is known to be consistent when the number of quantities and the length of the time series become large. As pointed out by Bai and Ng (2008), traditional methods usually overestimate the number of factors that are present in the data.
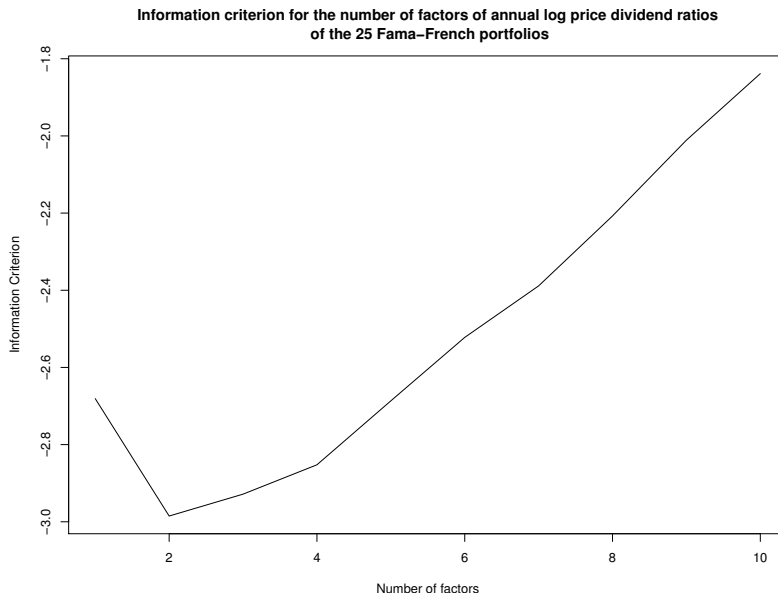
Figure 1: Information criterion as a function of the number of factors for the annual log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008.

22

**Scree plot of variances explained by the principal components of the
log price dividend ratios of the 25 Fama–French portfolios**

Figure 2: Variances explained by the first ten principal components of the annual log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008.

We carry out this procedure on the annual log P/D ratios of the 25 Fama French portfolios from 1943 (to account for the structural break). We find two significant factors in this series (as well as in the quarterly and monthly series of log P/D ratios – not reported).

We plot the information criterion as a function of the number of factors in figure 1 and the variances explained by the principal components in figure 2.

Using the same procedure, we find two factors in the first differences of the quarterly log P/D ratios of these portfolios. We also note that the plot of the variances explained by their principal components in figure 3 unambiguously points to a two factor structure.

We tabulate the rotations that relate the annual log P/D ratios of the 25 Fama-French portfolios to their first two principal components, denoted by $F_{1,2}$ with the subscript the principal component, in table 1. From the rotation matrices, we find that F_1 loads positively on all the portfolios and loads slightly more on the small stock portfolios. In contrast, F_2

Scree plot of variances explained by the principal components of the differences in log price dividend ratios of the 25 Fama–French portfolios
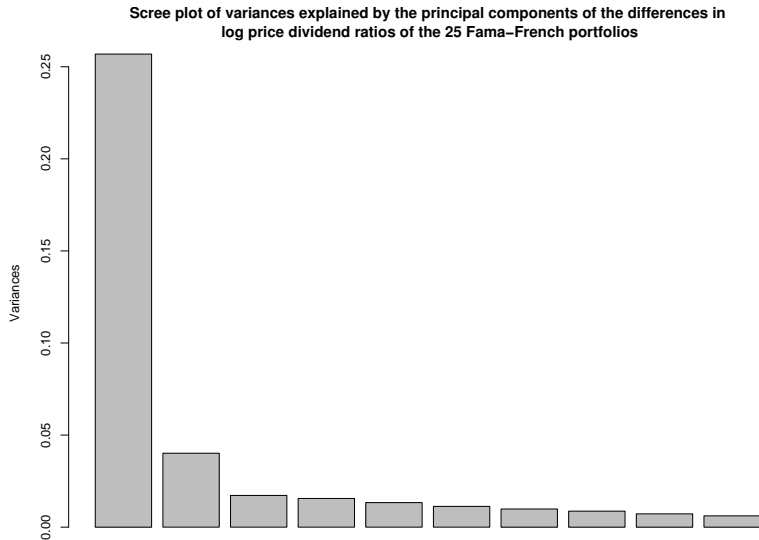
Figure 3: Since a factor structure for the log P/D ratios also implies a similar factor structure for the first differences in the log P/D ratios, we check that the first differences of the log P/D ratios also shows a two factor structure in the data. The above graph shows the variances explained by the first ten principal components of the first differences of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008. It shows that these first differences also exhibit the expected two factor structure.
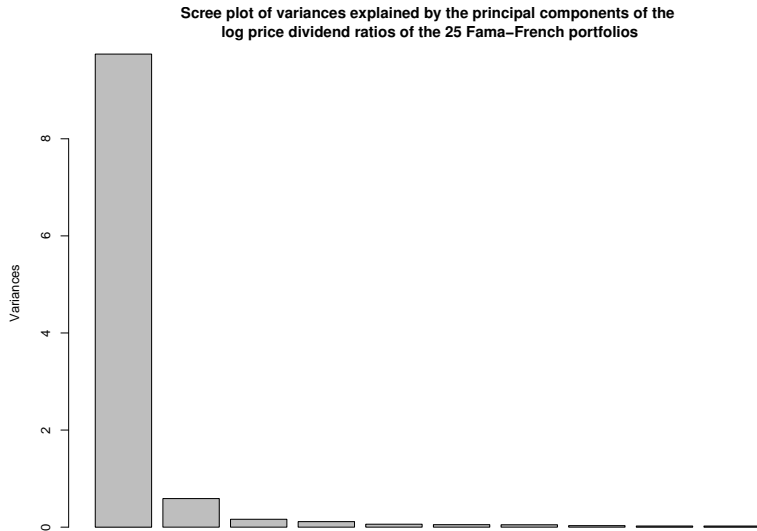
loads positively on large and value stocks and negatively on growth and small stocks. We thus expect F_2 to be closely related to the cross sectional differences among the portfolios.

We estimate the innovations of the two identified principal components as the OLS residuals obtained on regressing them on n lags of themselves, n being the smallest value for which they are serially uncorrelated at the 10% level according to both the Ljung-Box and Durbin-Watson tests. n is always found to be one for the annual data and sometimes two for the quarterly data (not reported). We denote these estimated innovations as IF_1 and IF_2.

24

		Rotation matrix for F_1				
		Growth				Value
		1	2	3	4	5
Small	1	0·356	0·267	0·206	0·194	0·169
	2	0·354	0·244	0·198	0·168	0·135
	3	0·314	0·210	0·176	0·152	0·135
	4	0·234	0·176	0·154	0·125	0·100
Large	5	0·148	0·116	0·104	0·114	0·116

		Rotation matrix for F_2				
		Growth				Value
		1	2	3	4	5
Small	1	−0·458	−0·140	−0·082	0·028	0·104
	2	−0·289	−0·055	0·089	0·194	0·211
	3	−0·177	0·025	0·185	0·231	0·221
	4	−0·064	0·091	0·170	0·302	0·243
Large	5	−0·005	0·077	0·186	0·275	0·315

Table 1: The rotation matrix that relates the log annual price dividend ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio to their first and second principal components. F_1 and F_2 represent the first and second principal components of the annual log P/D ratios over the period 1943-2008.

5.3 Principal Components & the Long Run Risk Factors

Since the X_i factors represent joint predictable components of consumption and dividend growth, a positive and significant coefficient should result on regressing future consumption and dividend growth against these factors. Similarly, since the V_j factors are components of the consumption growth volatility, regressing consumption growth volatility against them should also lead to a positive and significant coefficient. Since the principal component analysis only identifies affine transformations of the full set of long run risk factors, we can, in general, expect to find that the principal components will be related to both the X_i and V_j factors and that both regressions above will lead to significant coefficients given that the long run risk model holds. However, we find that only the volatility regression generates a significant coefficient for the first identified principal component F_1 and that only the future consumption and dividend growth regression generates a significant coefficient for the second identified

principal component F_2. This implies that the first identified principal component is naturally identifiable as an affine function of the only V factor and that the second identified principal component is naturally identifiable as an affine function of the only X factor.

We now examine the volatility regression in some detail. In order to construct a consumption volatility series, we estimate the innovations of quarterly consumption growth $\epsilon_{v,t}$ as the OLS residuals obtained on regressing it on n lags of itself, n being the smallest value for which they are uncorrelated at the 10% level according to both the Ljung-Box and Durbin-Watson tests. n is found to be three for this data series. Using these estimated innovations, the consumption volatility series is estimated as

$$v_t^n = \log \sum_{i=1}^{n} \frac{\epsilon_{v,\lfloor t+n/2-i \rfloor}^2}{n} \tag{32}$$

This methodology is standard and has been used in the context of long run risk models by Beeler and Campbell (2012).

The results of regressing v_t^{24}, v_t^{12} and v_t^6 on F_1^q and F_2^q, the factors constructed using quarterly log P/D ratios, are summarized in table 2. We also include the results of regressing de-trended v_t^{24}, v_t^{12} and v_t^6 on de-trended F_1^q as a robustness check as both consumption growth volatility and the first principal component have a pronounced time trend in the data as seen in figure 4. They show that F_1 is very closely related to consumption growth volatility with the R^2 of the 24 quarter volatility regression being as high as 81%. Even the R^2 for 6 quarter volatility regression, where the measurement error is likely to be high, is quite high at 47%. Further, the fact that the coefficients of F_1 in the various regressions are very similar to each other (i.e. for volatilities estimated over several horizons) provides strong evidence that the relation is robust. In contrast, there is no evidence at all that F_2 is related to consumption growth volatility. This result, when combined with the result, detailed below, that F_1 is unrelated to future consumption and dividend growth, leads to the conclusion that F_1 is an affine function of a V type factor. This conclusion follows because F_1 satisfies the conditions we have identified

for such a factor : it is an affine function of log P/D ratios, it tracks consumption growth

volatility and does not predict future consumption or dividend growth.[14]

Regression of consumption growth volatility on F_1 and F_2			
	F_1	F_2	R^2
24 quarter volatility	-0.213^{***} (0.027)	0.081 (0.094)	81.2%
12 quarter volatility	-0.234^{***} (0.052)	0.074 (0.190)	62.7%
6 quarter volatility	-0.235^{***} (0.062)	0.024 (0.209)	46.9%

Regression of de-trended consumption growth volatility on de-trended F_1		
	F_1	R^2
24 quarter volatility	-0.134^{***} (0.028)	32.2%
12 quarter volatility	-0.153^{**} (0.071)	13.6%
6 quarter volatility	-0.176^{*} (0.116)	9.5%

Table 2: Results of regressing volatility as defined in (32) against F_1 and F_2, the two significant principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1947-2008. The standard errors are Newey-West corrected with the required number of lags estimated using the procedure in (Newey and West 1994). The principal components of the quarterly log P/D ratios were used in these regressions. Use of the rotation matrix for the annual log P/D ratios to create quarterly versions of these principal components leads to very similar results.

We now examine the dividend and consumption growth regressions in detail. The results

of regressing annual real market dividend growth (i.e., growth of annual market dividends

deflated by the CPI) on the lagged values of F_1 and F_2 are summarized in table 3. We find,

from them, that F_2, but not F_1, predicts market dividend growth. This predictive ability is

weakly robust to lagging twice to account for time aggregation with the coefficient for F_2 being

significant at the 10% level.[15] (It should also be noted that time aggregation is generally not

[14]When analyzing the results obtained after de-trending, the reader should keep in mind that removing this time trend removes important information as neither quantity is expected to have a time trend on a theoretical basis. This is probably the reason for the lower R^2 obtained after the trends are removed. The most important aspect of these results, which is that the coefficient of de-trended F_1 is significant and has the same sign as without de-trending, still provides support to our conclusion.

[15]It is interesting that F_2, which weights the value portfolios more heavily, predicts future market dividend growth better than the log market P/D ratio (whose inability to predict dividend growth is well known (Cochrane 2005)). We hypothesize that this is because value stocks have a low duration which makes their P/D ratios depend more on dividend growth than on future expected excess returns.

12 quarter consumption growth volatility and the first principal component of the log price dividend ratios of the 25 Fama–French portfolios formed on the basis on size and B/M ratio
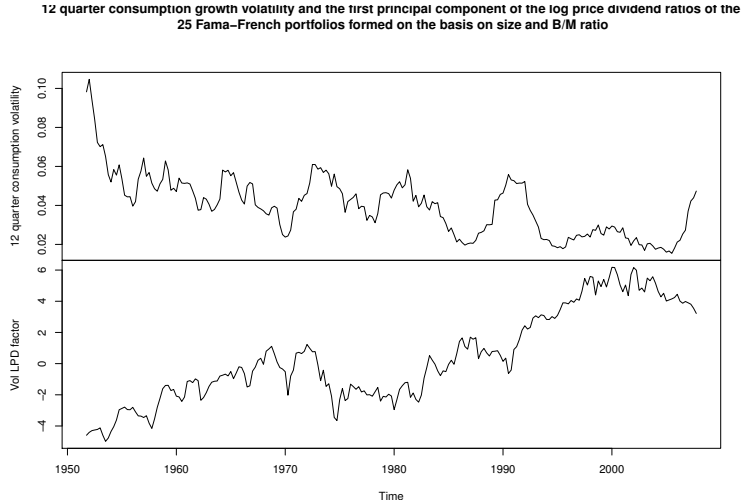
Figure 4: Plot of the 12 quarter consumption growth volatility and F_1, the first principal component of the log price dividend ratios of the 25 Fama-French portfolios formed on the basis of size and B/M ratio, over the period 1947-2008.

considered an issue with respect to dividend growth.) While our results are robust to time aggregation of dividend growth along this dimension, we acknowledge that such aggregation also leads to biases in our estimates of the price dividend ratios since we calculate them, as is conventional in the literature, using dividends aggregated on an annual basis in order to adjust for their pronounced seasonality. However, we also note that time aggregation of dividend growth, in contrast to time aggregation of consumption growth, is generally not considered a significant issue in the literature, that our model specification is annual rather than monthly and that our results are robust to the use of either of the conventional assumptions regarding the investment of dividends received during the year (the first being that such dividends are invested in nominal cash until the end of the year and the second being that they are invested in the asset itself until the end of the year[16]).

---

[16]The convention used in the presented calculations is equivalent to the assumption that dividends received during the year are consumed immediately and that the agent is completely indifferent to the timings of these dividends *during* the year.

The results of regressing annual real time consumption growth against the lagged values of F_1^a and F_2^a are also summarized in table 3. We find, from them, that F_2^a also predicts real time consumption growth as defined in the data section and that this predictive ability is robust to lagging twice to account for time aggregation. This is in accordance with the long run risk hypothesis that dividend and consumption growth share the same persistent component(s) X.[17] From the results in table 3, we conclude that F_2 can be identified as an affine function of a X type factor as it satisfies the essential properties of such factors : it is an affine function of log P/D ratios, predicts dividend and consumption growth but not consumption growth volatility.

We find that F_2 also satisfies another expected property of the X factor in many long run risk models. It has been pointed out by Bansal, Yaron, and Kiku (2007), Bansal, Dittmar, and Lundblad (2005), Bansal, Dittmar, and Kiku (2009), Da (2009) and others, the long run risk model implies that assets with higher sensitivity of predictable dividend growth to the long run risk factor X, which is measured by $\phi_{i,l}$ in our model, have higher expected excess returns. While this is not an implication of the long run risk model we use as we do not set $\pi_{i,l,x}$ to zero, our version still implies a positive relationship between $\phi_{i,l}$ and the expected excess return of asset l holding $\pi_{i,l,x}$ constant. Hence, if the long run risk model holds, we do find that the coefficients obtained on regressing portfolio dividend growths on F_X are significantly different from each other and are related to their excess returns in the expected manner. Specifically, we find that the F test strongly rejects the equality of the regression coefficients obtained on regressing real dividend growth of each of the 25 portfolios on the lagged value of F_2 at the $p < 0.001$ level and that these regression coefficients are higher for the portfolios of small and value stocks which have higher excess returns. We note that this result is not very surprising

---

[17]We note that while the use of this measure of consumption is not standard, it is more relevant for the current analysis as it better matches the information structure of the consumers in the economy. (It is also possible that real time data captures the sentiment of consumers as it reflects their current view of the state of the economy.)

with that on the lagged log market P/D ratio
(IPCD stands for the implicit personal consumption deflator)

	$F_{1,t}^a$	$F_{2,t}^a$	$\log(P/D)_{m,t}$	R^2
$\Delta d_{m,t+1}$ deflated by CPI	0.0004 (0.0031)	0.0317*** (0.0079)		16.0%
			0.003 (0.027)	0.0%
After one yr	-0.0005 (0.0032)	0.0134* (0.0089)		2.9%
$\Delta d_{m,t+1}$ deflated by IPCD	-0.0002 (0.0032)	0.0288*** (0.0079)		13.3%
			-0.002 (0.026)	0.0%
After one yr	-0.0012 (0.0033)	0.0126* (0.0088)		2.9%
$d_{m,t+3} - d_{m,t}$	0.0008 (0.0135)	0.0571** (0.0261)		11.6%
d_m formed by deflating with CPI			0.005 (0.107)	0.0%
$d_{m,t+3} - d_{m,t}$	-0.0012 (0.0132)	0.0543** (0.0247)		10.9%
d_m formed by deflating with IPCD			-0.011 (0.103)	0.0%

Regression of real time annual consumption growth on lagged F_1 and F_2

	$F_{1,t}$	$F_{2,t}$	R^2
Δc_{t+1}^{RT}	$-2 \times 10^{-4} (9 \times 10^{-4})$	0.0068* (0.0037)	17.4%
Δc_{t+2}^{RT}	$-5 \times 10^{-4} (7 \times 10^{-4})$	0.0054** (0.0021)	9.8%
$\Delta c_{t+1}^{RT} + \Delta c_{t+2}^{RT}$	$-5 \times 10^{-4} (1.5 \times 10^{-3})$	0.0123** (0.0050)	18.9%

Table 3: Results of regressing real annual market dividend growth (Δd_m where d_m is the log real annual market dividend) and real time consumption growth (Δc^{RT}) against lagged F_1 and F_2, the two significant principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio, and the lagged log market P/D ratio ($\log(P/D)_m$). The standard errors are Newey-West corrected with the required number of lags estimated using the procedure in (Newey and West 1994). The regressions using the log market P/D ratio are for the same time period as for the ones using F_1 and F_2 (1944-2008 for the one year dividend growth regressions, 1946-2008 for the three year dividend growth regressions). The one and two year real time consumption growth regressions are over the periods 1966-2008 and 1967-2008 respectively.

given the form of the rotation matrix relating F_2 to the log P/D ratios.[18] We do not test this relation explicitly as it is *not* an implication of the model we use.

Given the long run risk model, we also expect to find little if any cross-sectional variation of the sensitivity of dividend growth to the volatility factor. Empirically, we do find that this is largely the case with corresponding F statistic for F_1 (the principal component related to the volatility) being much lesser at 1.73 ($p = 0.015$). While this is marginally significant, it

---

[18]As in (Bansal, Dittmar, and Lundblad 2005), most of the individual coefficients are not significant but they are significantly *different* from each other.

is mostly because the regression coefficients for the portfolios corresponding to the smallest stocks being larger than the others.[19] Since the size premium is much less robust than the value premium, we see that this cross-sectional variation is not strongly related to expected excess returns.

Since the above regressions involve the whole sample and are subject to forward looking bias, we investigate whether the predictability implications that lead to the identification of the first two principal components as affine functions of the long run risk factors X and V hold out of sample in the section on asset pricing test results, and find that they indeed do so. We also show that look-ahead bias-free versions of the factors and their innovations estimated using the 25 Fama-French portfolios formed on the basis of size and book to market ratio successfully explain the expected returns of three sets of ten portfolios formed on the basis of short term reversal, long term reversal and the earnings to price ratio.[20] Hence, we see that our results are robust both to forward looking bias and are not an artifact of testing the model with the same portfolios as those that are used to estimate the factors.

5.4   Asset Pricing Tests

Since the dividends in this analysis have to be measured annually due to seasonality consider-ations, the asset pricing restrictions are only strictly correct at the annual time scale. Hence, we restrict ourselves to annual data in the following analysis.[21]

5.4.1   Cross Sectional Regressions

When we investigate the pricing relationship (15) using the 25 Fama-French portfolios, we find that it performs well. The cross sectional regression results are tabulated in table 4 and

---

[19]We used annual values for the above analysis in order to eliminate issues arising from dividend and CPI seasonality as well as to minimize the confounding effects that arise from overlapping regressions.

[20]They also explain the expected excess returns of the 25 Fama-French portfolios formed on the basis of size and book to market ratio.

[21]In unreported results, we find that the model also exhibits good performance at the quarterly time scale.

Results of the cross sectional regression

$$E[r_{i,t+\Delta t} - r_{f,t}] + \tfrac{1}{2}\mathrm{Var}[r_{i,t+\Delta t} - r_{f,t}] = \beta^2_{F_{-Vol}}\lambda_{F_{-Vol}} + \sum_{j=1}^{n+m}\beta_{IF_j}\lambda_{IF_j} + \beta_{\Delta c}\lambda_{\Delta c}$$

for the 25 Fama-French portfolios formed on the basis of size and book to market ratio

	Intercept	$\lambda_{F_{-Vol}}$	$\lambda_{IF_{-Vol}}$	λ_{IF_X}	$\lambda_{\Delta c}$	R^2
OLS	-0.027	-76.51	0.689	0.458	-0.0035	77.4%
	(-0.82)	(-2.79)	(2.96)	(3.97)	(-0.98)	(72.9%)
	-0.023		0.545	0.464	-0.0033	68.8%
	(-0.70)		(2.47)	(4.01)	(-0.93)	(64.4%)
	(-0.44)		(1.71)	(2.66)	(-0.62)	
WLS	-0.023		0.577	0.417	0.0000	
	(-0.48)		(1.94)	(2.66)	(0.01)	

Pricing errors $\times 100$

		Growth				Value
		1	2	3	4	5
Small	1	0.270	1.609	0.141	-0.874	-1.793
	2	1.899	-0.697	-1.430	0.065	-2.621
	3	-0.139	-0.145	-0.443	-0.245	-0.430
	4	-0.702	1.781	0.479	1.006	-0.503
Large	5	-0.518	1.507	-1.327	0.691	2.420

Table 4: Results of the two pass cross sectional regression, $E[r_{i,t+\Delta t} - r_{f,t}] + \tfrac{1}{2}\mathrm{Var}[r_{i,t+\Delta t} - r_{f,t}] = \beta^2_{F_{-Vol}}\lambda_{F_{-Vol}} + \sum_{j=1}^{n+m}\beta_{IF_j}\lambda_{IF_j} + \beta_{\Delta c}\lambda_{\Delta c}$, of the 25 Fama-French portfolios formed on the basis of size and book to market ratio, on lagged F_{-Vol} and concurrent IF_{-Vol} and IF_X over the period 1944-2008. F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors while IF_{-Vol} and IF_X are their innovations. For the OLS coefficients, the t values with and without the Shanken correction (Shanken 1992) (Shanken and Zhou 2007) are reported below the coefficient (the value without the correction is reported first) while for the WLS coefficients, only the t values with the correction are reported. The R^2 adjusted for the number of variables is reported below the unadjusted R^2.

plotted in figure 5. The OLS R^2 is quite high at 77.4%.[22] Further, the estimates of the cross-sectional regression intercept is insignificant, having an absolute t statistic value of 0.82. This, together with the insignificant intercepts and high R^2 values in the other cross sectional

---

[22]The pricing relationship (15) is also empirically supported at the quarterly frequency. While we do not report detailed results at this frequency for brevity, we note that the cross sectional regression intercept is not significantly different from zero and that the OLS R^2 is greater than 65%.

Cross sectional regression of the 25 Fama–French portfolios using consumption growth, the lagged value of the first principal component of their LPDs and innovations of the first two principal components of their LPDs
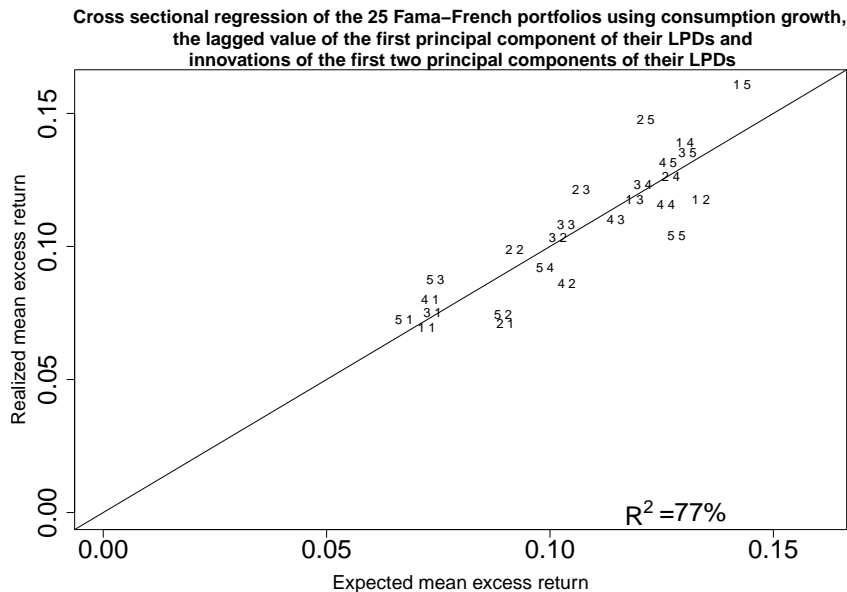
Figure 5: Results of the cross sectional regression of the 25 Fama-French portfolios formed on the basis of size and book to market ratio using IF_{-Vol}, IF_X and lagged values of F_{-Vol} over the period 1944-2008. F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors. IF_{-Vol} and IF_X are their innovations.

regressions detailed below provides support for our specification as the intercept must be zero for a correctly specified model, a fact emphasized by Jagannathan and Wang (2007).

Since we obtained the P/D ratio factors using the 25 Fama and French portfolios, it is necessary to check that these factors can price other portfolios as well. For that purpose we examine three sets of ten portfolios formed on the basis of short term reversal, long term reversal, and the E/P ratios. The cross sectional regression results are tabulated in table 5 and plotted in figure 6. We again find that the R^2 is high at 79.1% and the intercept is low at -1.6%/year and not statistically different from zero at conventional significance levels with the absolute value of the t statistic being only 0.71.

Results of the cross sectional regression for 30 portfolios

$$E[r_{i,t+\Delta t} - r_{f,t}] + \tfrac{1}{2}\text{Var}[r_{i,t+\Delta t} - r_{f,t}] = \beta_{F_{-Vol}}^2 \lambda_{F_{-Vol}} + \sum_{j=1}^{n+m} \beta_{IF_j} \lambda_{IF_j} + \beta_{\Delta c} \lambda_{\Delta c}$$

	Intercept	$\lambda_{F_{-Vol}}$	$\lambda_{IF_{-Vol}}$	λ_{IF_X}	$\lambda_{\Delta c}$	R^2
OLS	-0.016	-50.78	0.476	0.476	-0.0030	79.1%
	(-0.71)	(-1.65)	(2.51)	(4.36)	(-0.88)	(75.8%)
	-0.008		0.412	0.437	-0.0030	76.1%
	(-0.39)		(2.15)	(4.18)	(-0.88)	(73.4%)
	(-0.25)		(1.61)	(2.98)	(-0.61)	
WLS	-0.013		0.450	0.443	-0.0028	
	(-0.39)		(1.83)	(3.04)	(-0.57)	

Pricing errors $\times 100$

	bottom 1	2	3	4	5	6	7	8	9	top 10
Long term reversal	0·721	−0·915	−0·351	0·653	0·278	−0·636	−0·831	−0·003	1·700	0·404
Short term reversal	0·829	−2·109	−1·905	0·012	−0·059	1·404	0·967	0·322	−0·646	2·003
E/P ratio	−1·157	0·539	0·869	0·874	−0·250	−0·962	−1·073	0·106	−0·795	0·013

Table 5: Results of the cross sectional regression, $E[r_{i,t+\Delta t} - r_{f,t}] + \tfrac{1}{2}\text{Var}[r_{i,t+\Delta t} - r_{f,t}] = \beta_{F_{-Vol}}^2 \lambda_{F_{-Vol}} + \sum_{j=1}^{n+m} \beta_{IF_j} \lambda_{IF_j} + \beta_{\Delta c} \lambda_{\Delta c}$, of 30 portfolios (three sets of ten portfolios formed on the basis of long term reversal, short term reversal and the E/P ratio) on lagged F_{-Vol} and concurrent IF_{-Vol} and IF_X over the period 1952-2008. F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors. IF_{-Vol} and IF_X are their innovations. For the OLS coefficients, the t values with and without the Shanken correction (Shanken 1992) (Shanken and Zhou 2007) are reported below the coefficient (the value without the correction is reported first) while for the WLS coefficients, only the t values with the correction are reported. The R^2 adjusted for the number of variables is reported below the unadjusted R^2.

A common concern when using the cross sectional methodology is that the betas do not show sufficient cross sectional variation. However, we find that this is not the case. The F test for the hypothesis that the factor innovation betas are the same is strongly rejected for both factor innovations ($p < 0.001$). However, the analogous test for the lagged negative volatility factor cannot reject the hypothesis that these lagged volatility factor betas are different from each other even at the 10% level of significance. Hence, we also report the results excluding the lagged volatility factor betas in tables 4 and 5.

Cross sectional regression of 30 portfolios using consumption growth, the lagged value of the first principal component and the innovations of the first two principal components of the annual log price dividend ratios of the 25 Fama–French portfolios

Realized mean excess return

R^2 =79%
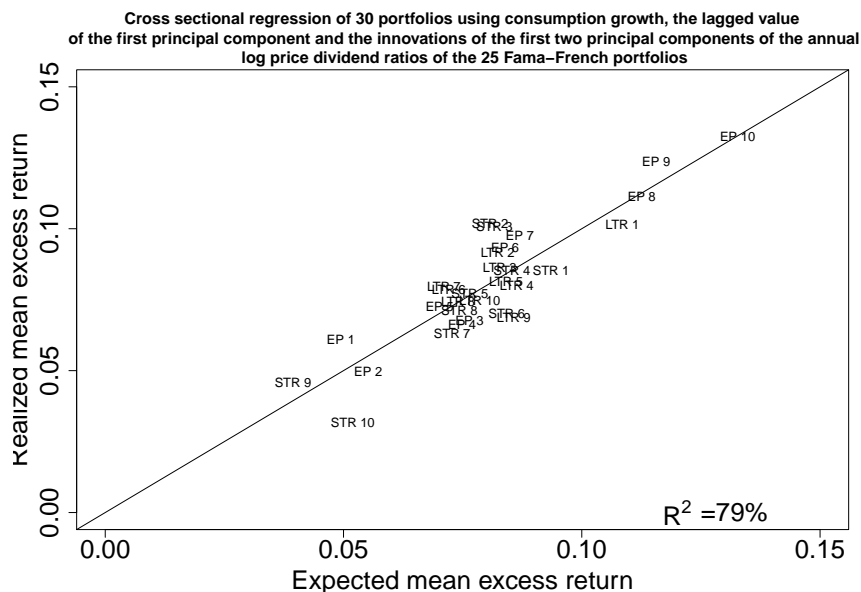
Expected mean excess return

Figure 6: Results of the cross sectional regression of 30 portfolios (three sets of ten portfolios formed on the basis of long term reversal, short term reversal and the E/P ratio) using lagged F_{-Vol}, IF_{-Vol} and IF_X over the period 1952-2008. F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors. IF_{-Vol} and IF_X are their innovations.

5.4.2 Robustness tests

Since the excess returns of the 25 Fama-French portfolios formed on the basis of size and book to market ratio have a strong factor structure, it is important to use robust test statistics to eliminate the problem of useless factors being identified as useful (a problem forcefully brought out by Lewellen, Nagel, and Shanken (2006), Kleibergen (2009) and Kleibergen (2010)). Hence, we test the above cross sectional regressions using the robust test statistics suggested by Kleibergen (2009) in appendix D to ensure that the factors here are not useless and that the hypothesis of factor pricing is not rejected. As shown in detail in this appendix, we find that both conditions are satisfied for our model specifications.

We also note that the number of time series observations in our analysis is small due to our low frequency data and that we find the betas of the assets to be significantly different from each other. As noted by Kan and Zhang (1999), these characteristics make it much less likely that a useless factor is spuriously found to be "useful" in a cross-sectional regression. Finally, we note that the cross-sectional regression intercept should be zero if the model is correctly specified, as emphasized by Jagannathan and Wang (2007), and that the intercept that we obtain for each of the cross-sectional regressions examined in the previous subsection is indeed close to zero and insignificantly different from it. In particular, the intercepts in all the cross sectional regressions that we investigate are insignificant and the absolute t statistic value is always less than 1.

As Ferson, Nallareddy, and Xie (2012) emphasize, it is important to check whether the results are robust to possible look ahead bias. We therefore investigate this later in this section. We find that our findings are not due to possible look ahead bias. In fact, we find that the bias free estimates of the log P/D factors and their innovations perform slightly better.

5.4.3 Price Dividend Ratio Factors vs Return Factors

The common practice in the empirical finance literature is to use excess return factors. We therefore examine the relation between them and the price dividend ratio factors.

We now interpret these results in terms of our earlier discussion relating price dividend ratio and return factors. Recall that it follows from (29) that the factor structure of returns is composed of three parts : (i) the factor structure of expected returns, (ii) the factor structure of innovations to the log price dividend ratios; and (iii) the factor structure of dividend growth innovations. As we show in appendix C, the latter can be chosen in a manner to make the log price dividend ratio factor structure consistent with the return factor structure.

Results of the constrained cross sectional regression for the 25 Fama-French portfolios formed on the basis of size and book to market ratio on the Fama-French factors, lagged F_{-Vol} and concurrent IF_{-Vol}, IF_X

$$E[r_{i,t+\Delta t}-r_{f,t}]+\tfrac{1}{2}\mathrm{Var}[r_{i,t+\Delta t}-r_{f,t}]=\beta^2_{F-Vol}\lambda_{F-Vol}+\sum_{j=1}^{n+m}\beta_{IF_j}\lambda_{IF_j}+\beta_{Mkt}\lambda_{Mkt}+\beta_{SMB}\lambda_{SMB}+\beta_{HML}\lambda_{HML}$$

	λ_{F-Vol}	λ_{IF-Vol}	λ_{IF_X}	λ_{Mkt}	λ_{SMB}	λ_{HML}
OLS	422.3	0.595	0.329	7.53	2.18	6.83
	(1.71)	(3.58)	(3.40)	(3.21)	(1.30)	(4.11)
OLS		0.624	0.344	7.72	2.46	6.77
		(3.76)	(3.72)	(3.33)	(1.47)	(4.07)
		(3.12)	(2.84)	(3.26)	(1.44)	(3.97)

Results of the constrained cross sectional regression for 30 portfolios formed on the basis of the earnings to price ratio, long term reversal and short term reversal on the Fama-French factors, IF_{-Vol}, IF_X and lagged F_{-Vol}

$$E[r_{i,t+\Delta t}-r_{f,t}]+\tfrac{1}{2}\mathrm{Var}[r_{i,t+\Delta t}-r_{f,t}]=\beta^2_{F-Vol}\lambda_{F-Vol}+\sum_{j=1}^{n+m}\beta_{IF_j}\lambda_{IF_j}+\beta_{Mkt}\lambda_{Mkt}+\beta_{SMB}\lambda_{SMB}+\beta_{HML}\lambda_{HML}$$

	λ_{F-Vol}	λ_{IF-Vol}	λ_{IF_X}	λ_{Mkt}	λ_{SMB}	λ_{HML}
OLS	-153.4	0.583	0.287	7.50	1.77	6.26
	(-1.32)	(3.78)	(3.76)	(2.98)	(0.62)	(2.74)
OLS		0.556	0.318	7.19	0.20	7.05
		(3.62)	(3.97)	(2.88)	(0.08)	(3.13)
		(3.18)	(3.25)	(2.83)	(0.06)	(2.70)

Table 6: Results of the constrained cross sectional regression of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008, and three sets of ten portfolios formed on the basis of earnings to price ratio, long term reversal and short term reversal over the period 1952-2008, on lagged F_{-Vol}, IF_{-Vol} and IF_X, and the Fama-French factors Mkt, SMB and HML. F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors. IF_{-Vol} and IF_X are their innovations. The t values with and without the Shanken correction (Shanken 1992) (Shanken and Zhou 2007) are reported below the coefficient (the value without the correction is reported first).

Another way to look at this, in the context of our model, is by using the following relation (see Appendix B, equation (91))

$$r_{l,t+1}-r_{f,t}=\sum_{j=1}^{m}\beta^e_{l,V_j}V_{j,t}+\beta_{l,c}(\Delta c_{t+1}-E_t[\Delta c_{t+1}])+\sum_{i=1}^{n}\beta_{l,Y_i}(X_{i,t+1}-E_t[X_{i,t+1}])$$
$$+\sum_{j=1}^{m}\beta_{l,Z_j}(V_{j,t+1}-E_t[V_{j,t+1}])+\epsilon_{l,t+1} \tag{33}$$

where the βs follow from the form of the dividend process.[23] It follows that, for our model, principal component analysis will recover the volatility factors, $V_{j,t}, j = 1, 2...m$, the innovation to the consumption growth of the representative agent, $(\Delta c_{t+1} - E_t[\Delta c_{t+1}])$,[24] innovations in the factors that drive expected consumption growth, $(X_{i,t+1} - E_t[X_{i,t+1}]), i = 1, 2, ..n$, innovations in the factors that drive volatility, $(V_{j,t+1} - E_t[V_{j,t+1}]), j = 1, 2..m$, and the factors in the dividend growth innovations $\epsilon_{l,t+1}$.

However, it is key to note that this principal component analysis will not recover the factors driving expected consumption growth, $X_{i,t}, i = 1, 2, \ldots, n$. Further, the factors driving expected consumption growth and volatility will explain only a small part of the variance in comparison with their innovations and $\epsilon_{l,t+1}$ may have some factor structure that may not be relevant for explaining the cross section of expected returns. This will lead to imprecise estimates of $V_{j,t}, j = 1, 2...m$. Hence, it will be difficult to evaluate the important implications of the long run risk models that $V_{j,t}, j = 1, 2...m$ should track consumption growth volatility over time and that $X_{i,t}, i = 1, 2...n$ should forecast consumption and dividend growth rates based on excess return factors alone.

Regression of real market dividend growth and real time consumption growth
on lagged F_X and Fama-French factors

	$Mkt - Rf_t$	SMB_t	HML_t	$F_{X,t}$	R^2
$\Delta d_{m,t+1}$	0.00053* (0.00031)	0.00004 (0.00066)	0.00047 (0.00051)	0.0266*** (0.0075)	19.0%
$d_{m,t+2} - d_{m,t}$	-0.00026 (0.00093)	-0.00002 (0.0012)	0.00102 (0.00068)	0.0472** (0.0179)	16.2%
$c_{t+2}^{RT} - c_t^{RT}$	0.00019 (0.00012)	0.00021 (0.00020)	0.00012 (0.00017)	0.00939** (0.0049)	25.6%

Table 7: Results of predictive regressions for market dividend (d_m) and real time consumption (c^{RT}) growth using the three Fama-French factors and the identified X type P/D factor, F_X, which is the second principal component of the log P/D ratios of the 25 Fama-French portfolios. The standard errors are Newey-West corrected with the required number of lags estimated using the procedure in (Newey and West 1994). The regressions for market dividend growth are over the period 1944-2008 while those for the real time consumption growth are over the period 1966-2008.

---

[23]It should be noted that this equation is specialized to our model while 29 is generally valid. For example, the $\beta_{l,c}$ is zero for the market in many standard long run risk models such as the ones studied by (Bansal and Yaron 2004) and (Constantinides and Ghosh 2008) and that an asset dependent constant will also be present if the volatility is an Ornstein-Uhlenbeck process.

[24]This will not be recovered in some long run risk models such as the one in (Bansal and Yaron 2004)

This point is empirically supported by the results documented in table 7. These results show that the estimated consumption/dividend growth factor using P/D ratios is able to predict dividend and real time consumption growth much better than the Fama-French factors. In fact, the regressions documented in the table show that the identified X type P/D factor is able to completely drive out the Fama-French factors when predicting these quantities. This factor is also far superior to just the use of the market price dividend ratio as documented in table 3.

In order to better understand the relationship between the Fama-French and P/D factors, we investigate the relation between the innovations of the log price dividend ratio factors and the Fama-French factors (which are excess return factors) as the latter have been proposed as proxies for future consumption growth by Parker and Julliard (2005) and consumption growth volatility by Boguth and Kuehn (2008). We summarize the results of regressing IF_{-Vol} and IF_X on the annual Fama-French factors in table 8. We find, from them, that IF_{-Vol} and IF_X can be approximately written as $Mkt + SMB$ and $Mkt + HML$ respectively. In other words, we find that, in the framework of this analysis, excess market returns are related to both consumption growth and consumption growth volatility, that SMB is related to consumption growth volatility and that HML is related to future consumption and dividend growth.

We can thus see that the Fama-French factors can be interpreted within the context of the long run risk model as capturing innovations to the long run risk factors with SMB capturing innovations to consumption growth volatility, HML capturing innovations to expected consumption growth and the excess market return capturing innovations in both these quantities. Interestingly, we find that (see table (6)) the Fama-French (return) factors do not drive out the long run risk factors and the long run risk factor innovations, which is consistent with these observations. Hence, the strong empirical support for the Fama and French three factor model does not necessarily mean that the data is inconsistent with long run risk models. The two models may be alternative representations of the same underlying phenomenon.

Regressions of IF_{-Vol} and IF_X on the Fama-French factors					
	Intercept	$R_m - R_f$	SMB	HML	R^2
IF_{-Vol}	-0.39^{***} (0.09)	3.95^{***} (0.42)	2.10^{***} (0.58)	0.50 (0.56)	68.8%
IF_X	-0.19^{***} (0.04)	1.42^{***} (0.18)	-0.22 (0.25)	1.42^{***} (0.25)	60.0%

Table 8: Results of regressing IF_{-Vol} and IF_X on the annual Fama-French factors. IF_{-Vol} and IF_X are the innovations of the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors. The regressions are over the period 1944-2008.

5.4.4 Addressing look ahead bias

To check the robustness of the results, we estimate the rotation matrices relating the log price dividend ratios of the portfolios to their first two principal components using only data from 1943 to 1975 and use them to construct out of sample version of the factors from 1976 to 2008. We find that these estimated out of sample factors, labelled F_1^{os} and F_2^{os}, track consumption growth volatility and predict market dividend and real time consumption growth in a manner similar to that documented for the in sample factors.

Regression of 24 quarter consumption growth volatility on F_1^{os} and F_2^{os}				
	Intercept	F_1^{os}	F_2^{os}	R^2
v_t^{24}	0.171^{***} (0.016)	-0.0050^{***} (0.0007)	0.0022 (0.0026)	74.1%

Table 9: Results of regressing real annual market dividend growth against lagged F_1^{os} and F_2^{os}, the out of sample estimates of the first and second price to dividend ratio factors. The regressions are over the period 1976-2008 while the factors' rotation matrices estimated using data from 1943-1975. The standard errors are Newey-West corrected with the required number of lags estimated using the procedure of (Newey and West 1994).

The results of regressing 24 quarter consumption growth volatility on the estimated out of sample factors, summarized in table 9, show that the relation found in the paper is robust. Specifically, consumption growth volatility is found to be very significantly negatively related to the first out of sample factor F_1^{os} and to be unrelated to the second out of sample factor F_2^{os}.

Regression of market dividend growth on F_1^{os} and F_2^{os} and the
log market price dividend ratio

	F_1^{os}	F_2^{os}	$\log(P/D)_m$	R^2
1 yr. Mkt div. growth	-0.0066 (0.0055)	0.0491*** (0.0183)		20.8%
			0.012 (0.036)	0.8%
3 yr. Mkt div. growth	0.0026 (0.0263)	0.0593 (0.0453)		13.4%
			0.066 (0.138)	5.7%

Regression of real time annual consumption growth on lagged values of F_1^{os} and F_2^{os}

	F_1^{os}	F_2^{os}	R^2
Δc_{t+1}^{RT}	$4.1 \times 10^{-4}(0.0010)$	0.0063** (0.0031)	13.9%
Δc_{t+2}^{RT}	$5.5 \times 10^{-4}(6.8 \times 10^{-4})$	0.0045** (0.0016)	5.8%
$\Delta c_{t+1}^{RT} + \Delta c_{t+2}^{RT}$	$7.6 \times 10^{-4}(1.5 \times 10^{-3})$	0.0123** (0.0050)	18.9%

Table 10: Results of regressing real annual market dividend growth and real time consumption growth (Δc^{RT}) against lagged F_1^{os} and F_2^{os}, the out of sample estimates of the first and second price to dividend ratio factors. The regressions are over the period 1976-2008 while the factors' rotation matrices estimated using data from 1943-1975. The standard errors are Newey-West corrected with the required number of lags estimated using the procedure of (Newey and West 1994). The regressions using the log market price dividend ratio use data from 1976 onwards in order to be consistent with the others.

The predictability of real time consumption and market dividend growth using the out of sample factors are summarized in table 10. We can see that only the second factor is relevant in predicting real time consumption growth and market dividend growth. The result for the three year market dividend growth seems marginal but that is because the number of data points is much smaller and the R^2 of the regression is still found to be quite high.

To further address the critique of Ferson, Nallareddy, and Xie (2012), we also undertake a complete re-analysis of the model where we remove the possibility of look-ahead bias. We do so by re-calculating the rotation matrix for the principal components, and therefore the factors and their innovations, every year with these components and their innovations being calculated based on the rotation matrix estimated with data up to the previous year. We ensure that a minimum of 25 years of data is used for the calculation of each rotation matrix. We denote these no look ahead bias version of the factors as F_1^{nlb} and F_2^{nlb}.

Results of the cross sectional regression

$$E[r_{i,t+\Delta t} - r_{f,t}] + \tfrac{1}{2}\mathrm{Var}[r_{i,t+\Delta t} - r_{f,t}] = \beta_{F-Vol}^2 \lambda_{F-Vol} + \sum_{j=1}^{n+m} \beta_{IF_j}\lambda_{IF_j} + \beta_{\Delta c}\lambda_{\Delta c}$$

using a look-ahead bias free approach for 30 portfolios based on long and short term reversal and the E/P ratio

	Intercept	$\lambda_{F^{nlb}_{-Vol}}$	$\lambda_{IF^{nlb}_{-Vol}}$	$\lambda_{IF^{nlb}_X}$	$\lambda_{\Delta c}$	R^2
OLS	0.0005	-117.7	0.339	0.410	-0.0033	70.0%
	(0.02)	(-1.97)	(1.39)	(2.71)	(-0.96)	(68.9%)
OLS	0.014		0.223	0.381	-0.0031	61.9%
	(0.49)		(0.89)	(2.54)	(-0.90)	(60.5%)
	(0.33)		(0.66)	(1.82)	(-0.66)	
WLS	0.010		0.273	0.362	-0.0021	
	(0.26)		(0.92)	(1.97)	(-0.53)	

Pricing errors $\times 100$

	bottom 1	2	3	4	5	6	7	8	9	top 10
Long term reversal	0·748	−1·402	−0·424	1·416	−0·233	−0·792	−1·201	−0·823	0·856	−0·446
Short term reversal	−0·106	−2·243	−1·097	−1·392	−0·343	1·654	1·311	0·425	0·851	2·825
E/P ratio	0·583	1·064	0·043	0·554	−0·695	0·178	−1·075	−0·502	0·028	0·240

Table 11: Results of the cross sectional regression, including pricing errors, of 30 portfolios (three sets of ten portfolios formed on the basis of long term reversal, short term reversal and the E/P ratio) on lagged F^{nlb}_{-Vol} and concurrent IF^{nlb}_{-Vol} and IF^{nlb}_X from 1968-2008. F^{nlb}_{-Vol} and F^{nlb}_X are the negative volatility and consumption/dividend growth factors estimated in a manner free of look-ahead bias and IF^{nlb}_{-Vol} and IF^{nlb}_X are their innovations. For the OLS coefficients, the t values with and without the Shanken correction (Shanken 1992) (Shanken and Zhou 2007) are reported below the coefficient (the value without the correction is reported first) while for the WLS coefficients, only the t values with the correction are reported. The R^2 adjusted for the number of variables is reported below the unadjusted R^2.

Such a calculation results in reasonable look-ahead bias estimates of the innovations of the factors as the innovations are estimated in a manner which only makes use of a single rotation matrix which is calculable by the agent at that time.[25] Hence, this version of the robustness check is most important for establishing the robustness of the cross sectional regressions to look-ahead bias as the cross sectional regression results are mostly dependent on the factor innovations.[26]

---

[25]The estimation of the factors themselves is, however, likely to be quite noisy using this approach as different rotation matrices are used for each year. This why the predictability analysis in tables 9 and 10 was carried out with look ahead bias free factors estimated with a *single* rotation matrix.

[26]We also, however, find that the results of this robustness check for the predictability implications is largely supportive of our hypothesis and the results are available upon request.

Cross sectional regression of 30 portfolios using consumption growth, the lagged value of the first principal component and the innovations of the first two principal components of the annual log price dividend ratios of the 25 Fama–French portfolios with no look ahead bias
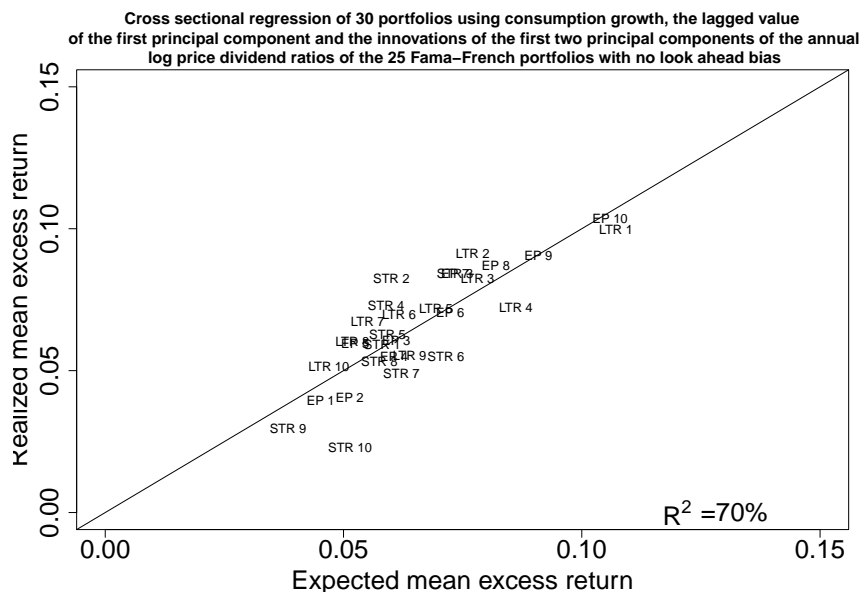
Figure 7: Results of the cross sectional regression of 30 portfolios (three sets of ten portfolios formed on the basis of long term reversal, short term reversal and the E/P ratio) using lagged F_{-Vol}^{nlb} and concurrent IF_{-Vol}^{nlb}, IF_X^{nlb} and Δc innovations. F_{-Vol}^{nlb} and F_X^{nlb} are the negative volatility and consumption/dividend growth factors constructed in a look-ahead bias free manner and IF_{-Vol}^{nlb} and IF_X^{nlb} are their innovations. The cross sectional regression is done over the period 1968-2008.

For the sake of brevity, we only report the results for the cross-sectional regression of the thirty portfolios based on long term reversal, short term reversal and the earnings to price ratio on look-ahead bias free versions of the factors and their innovations as it also removes the concern that the factors are estimated with the same portfolios as are being tested. These results are summarized in table 11 and are plotted in figure 7. The results are highly supportive for the model as the R^2 is high and that the intercept of 0.0005 is essentially zero. The latter is important as it is known from the study of Lewellen, Nagel, and Shanken (2006) that the R^2 by itself is not very significant and that a close to zero intercept with a high R^2 provides

significantly more evidence for the model than just a high R^2 (a fact also emphasized by Jagannathan and Wang (2007)).[27]

5.4.5 GMM Tests

While the cross sectional regression methodology above provides a nice, intuitive way of under-standing the importance of the different variables in the stochastic discount factor, it can only handle linear relationships and needs relatively restrictive assumptions for accurate results. Since the exact Euler equation restrictions

$$E_t[M_{t+1}R_{t+1}^e] = 0 \tag{34}$$

are nonlinear in nature, we now use GMM to ensure that the above results are robust. The results of the GMM estimation of (34) using the 25 Fama-French portfolios sorted on the basis of size and book to market ratio are summarized in table 12. The corresponding results for the three sets of ten portfolios sorted on the basis of long term reversal, short term reversal and E/P ratio are summarized in table 13 .

Since excess returns are used in these tests, the mean of the stochastic discount factor must be accounted for by adding an additional moment condition (as pointed out by Kan and Robotti (2008)). We report the results obtained after adding this moment condition in tables 14 and 15 respectively. For the former, we only report the results obtained using the identity weighting matrix as the estimated optimal weighting matrix is almost singular. These results are found to be closer to those obtained using the cross sectional regression approach. In particular, Γ_1 is found to be fairly large and significant and α_v is found to be negative for the 25 Fama-French portfolios, though not significantly so.

---

[27]While consumption growth was included in this cross sectional regression in order to illustrate the relative unimportance of consumption growth, it is clear from the almost zero value of $\lambda_{\Delta c}$ that it's addition or removal has almost no impact on the results.

Identity weighting matrix		Optimal weighting matrix	
Coefficient	Estimate (Std. Err.)	Coefficient	Estimate (Std. Err.)
$\Gamma_1\left(\frac{X}{F_X}\right)$	0.463 (0.683)	$\Gamma_1\left(\frac{X}{F_X}\right)$	0.433 (0.282)
$\Gamma_2\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	0.056 (0.107)	$\Gamma_2\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	0.036 (0.069)
$\alpha_{IF_X}=\alpha_x\left(\frac{X}{F_X}\right)$	1.734 (0.742)	$\alpha_{IF_X}=\alpha_x\left(\frac{X}{F_X}\right)$	1.559 (0.448)
$\alpha_{IF_{-Vol}}=\alpha_v\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	0.114 (0.236)	$\alpha_{IF_{-Vol}}=\alpha_v\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	0.106 (0.149)
Dist statistic	0.1086	J statistic	23.97 (p=0.29)

Table 12: GMM test of the Euler equation restrictions $E[M_{t+1}R^e_{t+1}]=0$ for the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008 without performing the correction for the effect of the unspecified mean of the stochastic discount factor suggested by (Kan and Robotti 2008). For this, we make use of the expression for the log stochastic discount factor $\log M_{t+\Delta t}=\Delta t(\Gamma_0+\Gamma_1 X+\Gamma_2(V-\bar{V})-\alpha_x(X_{t+\Delta t}-E_t[X_{t+\Delta t}])-\alpha_v(V_{t+\Delta t}-E_t[V_{t+\Delta t}])-\alpha_c(c_{t+\Delta t}-E_t[c_{t+\Delta t}]))$, where X and V are the underlying long run risk state variables, and drop the last term since it is known to be unimportant (we have verified that it's inclusion does not change the results). F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors which are affinely related to X and V. IF_{-Vol} and IF_X are their innovations. The affine relationship between X and F_X and V and F_{-Vol} together with the zero mean of X imply that $\frac{X}{F_X}$ and $\frac{V-\bar{V}}{F_{-Vol}}$, where $\bar{V}$ is the mean of V, are the linear coefficients of the affine relationships. Hence, the identified coefficients of F_{-Vol} and F_X and their innovations in the GMM estimation are related to the corresponding coefficients of the state variables in the manner indicated.

α_x, the market price of risk for shocks to expected consumption growth, is notably highly significantly positive in all of the GMM estimations with estimates of it's scaled value ranging from 1.56 to 2.90. Since α_x is directly related to the coefficient of relative risk aversion γ, we can use these estimates to obtain an estimate for γ.

We note that the identification of the two factors in this study also enables the determination of the relative importance of cash flow and discount rate risks for cross-sectional returns in the context of long run risk models. This is because the rate at which future equity cash flows are discounted (the equity risk premium) is determined by the consumption growth volatility in these models as shown by Bansal and Yaron (2004) and others, which in turn means that the first factor proxies for discount rate risk and that the second proxies for cash flow risk. The results of the analysis using the innovations of the two factors indicate that cash flow risk is cross-sectionally more important than discount rate risk. This result is robust to the inclusion

Identity weighting matrix		Optimal weighting matrix	
Coefficient	Estimate (Std. Err.)	Coefficient	Estimate (Std. Err.)
$\Gamma_1 \left(\frac{X}{F_X} \right)$	0.329 (0.480)	$\Gamma_1 \left(\frac{X}{F_X} \right)$	0.254 (0.295)
$\Gamma_2 \left(\frac{V - \bar{V}}{F_{-Vol}} \right)$	-0.058 (0.117)	$\Gamma_2 \left(\frac{V - \bar{V}}{F_{-Vol}} \right)$	-0.075 (0.068)
$\alpha_{IF_X} = \alpha_x \left(\frac{X}{F_X} \right)$	2.139 (0.884)	$\alpha_{IF_X} = \alpha_x \left(\frac{X}{F_X} \right)$	1.758 (0.524)
$\alpha_{IF_{-Vol}} = \alpha_v \left(\frac{V - \bar{V}}{F_{-Vol}} \right)$	-0.050 (0.286)	$\alpha_{IF_{-Vol}} = \alpha_v \left(\frac{V - \bar{V}}{F_{-Vol}} \right)$	0.004 (0.181)
Dist statistic	0.0603	J statistic	25.43 (p=0.49)

Table 13: GMM test of the Euler equation restrictions $E[M_{t+1} R_{t+1}^e] = 0$ for the three sets of ten portfolios formed on the basis of the earnings to price ratio, long term reversal and short term reversal over the period 1952-2008 without performing the correction for the effect of the unspecified mean of the stochastic discount factor suggested by (Kan and Robotti 2008). See caption for table 12 for details regarding the notation.

of the lagged factors, as is seen from the GMM results summarized below. This study thus underlines the importance of cash flow risk and contributes to the recent strand of literature that demonstrates that it can explain a large proportion of the cross-sectional return variation (Campbell and Vuolteenaho 2004) (Bansal, Dittmar, and Lundblad 2005) (Cohen, Polk, and Vuolteenaho 2008) (Campbell, Polk, and Vuolteenaho 2009) (Da and Warachka 2009).

5.5 Relative Risk Aversion

Since it is largely F_X that predicts future dividend and consumption growth, we can, as pointed out, among others, by Hansen, Heaton, and Li (2008) and Kaltenbrunner and Lochstoer (2010), use the GMM estimate for the coefficient of scaled innovation of X (i.e. IF_X) in the stochastic discount factor, α_{IF_X}, to make an estimate of the preference for the early resolution of uncertainty, i.e. $\gamma - 1/\psi$. Since the value of the elasticity of intertemporal substitution (EIS), ψ, has to be large in long run risk models in order for them to be consistent with the low volatility of the real risk free rate, this value also provides an estimate of the relative risk aversion γ since $\gamma - < 1/\psi \approx \gamma$ if $\psi > 1$.[28]

---

[28]We also note that a large value of ψ is not only strongly suggested by the low volatility of the real risk free rate but also by the analysis of household survey data by Vissing-Jørgensen and Attanasio (2003) and by the study of an elegant natural experiment by Kapoor and Ravi (2010).

Identity weighting matrix	
Coefficient	Estimate (Std. Err.)
$\Gamma_1\left(\frac{X}{F_X}\right)$	2.117 (0.542)
$\Gamma_2\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	0.115 (0.212)
$\alpha_{IF_X}=\alpha_x\left(\frac{X}{F_X}\right)$	2.314 (0.297)
$\alpha_{IF_{-Vol}}=\alpha_v\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	-0.091 (0.122)
Dist statistic	0.334

Table 14: GMM test of the Euler equation restrictions $E[M_{t+1}R^e_{t+1}]=0$ for the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008, together with the correction for the biases introduced due to the unspecified mean of the stochastic discount factor suggested by (Kan and Robotti 2008). For the computation, we make use of the expression for the log stochastic discount factor $\log M_{t+\Delta t}=\Delta t(\Gamma_0+\Gamma_1 X+\Gamma_2(V-\bar{V})-\alpha_x(X_{t+\Delta t}-E_t[X_{t+\Delta t}])-\alpha_v(V_{t+\Delta t}-E_t[V_{t+\Delta t}])-\alpha_c(c_{t+\Delta t}-E_t[c_{t+\Delta t}]))$, where X and V are the underlying long run risk state variables, and drop the last term since it is known to be unimportant (we have verified that it's inclusion does not change the results). F_{-Vol} and F_X are the first two principal components of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1943-2008 which, as we show, can be interpreted as negative volatility and consumption/dividend growth factors which are affinely related to X and V. IF_{-Vol} and IF_X are their innovations. The affine relationship between X and F_X and V and F_{-Vol} together with the zero mean of X imply that $\frac{X}{F_X}$ and $\frac{V-\bar{V}}{F_{-Vol}}$, where $\bar{V}$ is the mean of V, are the linear coefficients of the affine relationships. Hence, the identified coefficients of F_{-Vol} and F_X and their innovations in the GMM estimation are related to the corresponding coefficients of the state variables in the manner indicated. We carry out this test only with the identity matrix as the optimal weighting matrix is very close to singular. Note that the Dist statistic is *not* comparable with that in table 12 as the means of the stochastic discount factors are different.

We obtain this estimate by first noting the relation (which follows from the identification $F_X \propto X$ and the fact that there is only one identified X component)

$$\begin{aligned}\alpha_{IF_X} &= \frac{\gamma-1/\psi}{1-\nu_1(1-\alpha\Delta t)}\left(\frac{\epsilon_t}{IF_{X,t}}\right) \\ &= \frac{\gamma-1/\psi}{1-\nu_1(1-\alpha\Delta t)}\left(\frac{X}{F_X}\right)\end{aligned} \tag{35}$$

where $IF_{X,t}$ and $\epsilon_t=\varphi_x\delta_x\sqrt{V}(Y_t-Y_{t-1})$ stand for the innovations of the second principal component and X respectively (note that $IF_{X,t}$ and ϵ_t are proportional to each other here so that their ratio is still independent of t). An estimate for $\frac{X}{F_X}$ can be obtained from the results of the regression in table 3. Using the relation (1), it is not difficult to see that coefficient

Identity weighting matrix		Optimal weighting matrix	
Coefficient	Estimate (Std. Err.)	Coefficient	Estimate (Std. Err.)
$\Gamma_1\left(\frac{X}{F_X}\right)$	1.164 (0.761)	$\Gamma_1\left(\frac{X}{F_X}\right)$	1.165 (0.284)
$\Gamma_2\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	-0.002 (0.120)	$\Gamma_2\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	-0.008 (0.078)
$\alpha_{IF_X}=\alpha_x\left(\frac{X}{F_X}\right)$	2.882 (1.133)	$\alpha_{IF_X}=\alpha_x\left(\frac{X}{F_X}\right)$	2.899 (0.269)
$\alpha_{IF_{-Vol}}=\alpha_v\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	-0.241 (0.301)	$\alpha_{IF_{-Vol}}=\alpha_v\left(\frac{V-\bar{V}}{F_{-Vol}}\right)$	-0.257 (0.091)
Dist statistic	0.228	J statistic	21.37 (p=0.72)

Table 15: GMM test of the Euler equation restrictions $E[M_{t+1}R^e_{t+1}]=0$ for the three sets of ten portfolios formed on the basis of the earnings to price ratio, long term reversal and short term reversal over the period 1952-2008, together with a correction for the biases introduced by the unspecified mean of the stochastic discount factor suggested by (Kan and Robotti 2008). See the caption for table 14 for details regarding the notation. Note that the Dist statistic is *not* comparable with that in table 13 as the means of the stochastic discount factors are different.

obtained when regressing $\Delta c_{t+1} + \Delta c_{t+2}$ on X_t is given by $\sum_{i=0}^{1}(1-\alpha\Delta t)^i$. Hence, the coefficient obtained when regressing $\Delta c_{t+1} + \Delta c_{t+2}$ on F_X is

$$\sum_{i=0}^{1}(1-\alpha\Delta t)^i \frac{F_X}{X} \tag{36}$$

By noting that the persistence of the X process is the same as that of F_X since they are proportional, we find α to be about 0.15 on the annual time scale. Using this value and the regression coefficient of 0.0123 obtained in table 3, we find that $\frac{F_X}{X} \approx 150$. This value, together with the annual estimate of $0.997^{12} \approx 0.97$ for ν_1 from Bansal and Yaron (2004) and the GMM estimates of 1.56 to 2.89 for the market price of risk of innovations to F_X, give an estimate of between 40 and 75 for $\gamma - 1/\psi$ (or equivalently γ since $\gamma \gg 1/\psi$). While high, this estimate is similar to the value of 60 obtained by Chen, Favilukis, and Ludvigson (2007).

The leverage of market dividend growth on long term consumption growth can be similarly estimated from the results in table 3. It is found to be about 3.3 when three year market dividend growth is used in the analysis. This value is remarkably similar to that proposed by Bansal and Yaron (2004).

It should be noted that the main reason for the risk aversion estimate to be much higher than that proposed by Bansal and Yaron (2004) is that the volatility of consumption growth after the structural break (1.85×10^{-4}) is much lower than for the entire period for which data is available (4.92×10^{-4}). If we scale the relative risk aversion value estimates that we obtain by the ratio of these volatilities, we find that it is very similar to the value of 16 obtained by Bansal, Yaron, and Kiku (2007). Hence, it is possible that a long run risk model which accounts for structural breaks or regime shifts in the parameters will require a much lower relative risk aversion to explain asset prices as such a model can have a much higher unconditional volatility of consumption growth and still be consistent with the data. We also note that the standard errors for our estimate are large and we cannot rule out that the relative risk aversion value is below 10 at the 1% level of significance.

6   Conclusion

In this paper, we show that typical long run risk models, including those of Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Zhou and Zhu (2009), imply that the log P/D ratios of financial assets have a strict factor structure when the intertemporal budget constraint of the marginal investor can be well approximated by the loglinear method of Campbell and Shiller (1988). Further, we demonstrate that these factors must be related to aggregate consumption growth and consumption growth volatility when there is a representative agent. When we restrict attention to the post-1942 data so as to account for the structural break documented by Marakani (2009), we find that the log P/D ratios of the 25 Fama-French portfolios have two significant factors, one of which is related to aggregate consumption growth volatility and the other to future dividend and real time aggregate consumption growth.

These factors and their innovations do a reasonably good job of explaining the cross section of returns of not only the 25 portfolios from which they were formed but also three sets of ten portfolios based on long term reversal, short term reversal and the earnings to price ratio.

The coefficients obtained from the cross sectional regressions are statistically and economically significant and have the right sign, and the zero beta rate is economically small and not significantly different from zero.

We address the critique of Ferson, Nallareddy, and Xie (2012) that the original results of Bansal, Yaron, and Kiku (2007) are not robust to look ahead bias. Our results go through even when we use versions of the factors which are free of look-ahead bias. This difference is due to the fact that we use a different method for estimating the unobserved factors that is robust to the presence of measurement errors in consumption and the real risk free rate. We also address one of the important critiques of the long run risk model raised by Beeler and Campbell (2012). Beeler and Campbell (2012) point out that the log market price dividend ratio does not predict dividend or consumption growth. We show that a price dividend ratio factor does predict both dividend growth a real time measure of consumption growth. We argue that this does not show up in the log market price dividend ratio as the market price dividend ratio weights most strongly on the volatility price dividend ratio factor.

Our findings link the classical commonly used linear return factor models in the finance literature with the more recent long run risk models. Our analysis suggests that they can be viewed as alternative representations of the same underlying phenomenon. The advantage of our approach lies in the fact that it additionally enables the estimation of all the underlying state variables of the long run risk model which is not possible with return factor models alone. More specifically, we show that the return factors do not include the expected consumption/dividend growth factor of long run risk models even though they are estimable using the P/D factors. This is corroborated empirically as our estimated consumption/dividend growth factor drives out the Fama-French factors in predictive regressions of dividend and real time consumption growth.

The link we establish between linear return factor models and long run risk models also shows that the strong empirical support for the Fama and French three factor model does not

necessarily mean that the data are inconsistent with long run risk models. The two models may be alternative representations of the same underlying phenomenon. The advantage of using the proposed methodology is that it further enables us to estimate the factors of the long run risk model themselves and hence test implications of the long run risk model which cannot be investigated with return factors alone. The link also enables us to interpret the Fama-French factors within the context of the long run risk model.

As Beeler and Campbell (2012) point out, long run risk models imply counter-factually high predictability of long term aggregate consumption growth, long term dividend growth and future market volatility by the market price-dividend ratio. In this paper, we address the first two issues by showing that a log P/D factor does in fact predict long term dividend growth and real time consumption growth. While we do not consider market volatility in this paper, we do find, in unreported results, some indicative evidence that the log P/D factors also predict some aspects of market volatility.

Beeler and Campbell (2012) also point out that long run risk models imply counter-intuitively high or infinite prices for real risk free consol bonds. This weakness of the long run risk model (and many other asset pricing models) is related to the fact that the variance of the permanent and transitory components of it's stochastic discount factor are inconsistent with the data as pointed out by Bakshi and Chabi-Yo (2012). Addressing this issue may require considering alternative stochastic processes for modeling long run risk including trend stationary models and regime switching models.

References

Adrian, T., E. Etula, and T. Muir, 2011, "Financial intermediaries and the cross-section of asset returns," *Staff Reports*, 464.

Albuquerque, R., M. Eichenbaum, and S. Rebelo, 2012, "Understanding the equity-premium

and correlation puzzles," Private communication.

Bai, J., and S. Ng, 2008, "Large Dimensional Factor Analysis," *Foundations and Trends in Econometrics*, 3(2), 89 – 168.

Bakshi, G., and F. Chabi-Yo, 2012, "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors," *Journal of Financial Economics*, 105, 191–208.

Bansal, R., R. F. Dittmar, and D. Kiku, 2009, "Cointegration and consumption risks in asset returns," *Review of Financial Studies*, 22(3), 1343–1375.

Bansal, R., R. F. Dittmar, and C. Lundblad, 2005, "Consumption, Dividends, and the Cross-section of Equity Returns," *Journal of Finance*, 60, 1639–1672.

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

Bansal, R., A. Yaron, and D. Kiku, 2007, "Risks for the Long Run : Estimation and Inference," Working paper.

———— , 2012, "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

Barro, R., 2006, "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823–866.

Beeler, J., and J. Y. Campbell, 2012, "The Long-Run Risks Model and Aggregate Asset Prices : An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bekaert, G., and E. C. Engstrom, 2010, "Asset Return Dynamics under Bad Environment-Good Environment Fundamentals," *SSRN eLibrary*.

Boguth, O., and L. Kuehn, 2008, "Consumption Volatility Risk," *SSRN eLibrary*.

Campbell, J., and J. Cochrane, 1999, "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of political Economy*, 107(2), 205–251.

Campbell, J., C. Polk, and T. Vuolteenaho, 2009, "Growth or glamour? Fundamentals and systematic risk in stock returns," *Review of Financial Studies*.

Campbell, J. Y., and R. J. Shiller, 1988, "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1, 195–227.

Campbell, J. Y., and T. Vuolteenaho, 2004, "Bad Beta, Good Beta," *The American Economic Review*, 94(5), 1249–1275.

Chen, X., J. Favilukis, and S. C. Ludvigson, 2007, "An Estimation of Economic Models with Recursive Preferences," FMG Discussion Papers dp603,, Financial Markets Group.

Cochrane, J. H., 2005, *Asset Pricing*. Princeton University Press.

Cohen, R., C. Polk, and T. Vuolteenaho, 2008, "The price is almost right," *Journal of Finance*, 64, 2739–2782.

Connor, G., and R. A. Korajczyk, 1986, "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis," *Journal of Financial Economics*, 15, 373–394.

——— , 2009, "Factor Models of Asset Returns," *SSRN eLibrary*.

Constantinides, G. M., and A. Ghosh, 2008, "Asset Pricing Tests with Long Run Risks in Consumption Growth," *SSRN eLibrary*.

Croushore, D., 2010, "Philadelphia Fed Forecasting Surveys : Their Value for Research," *Business Review*, Q3, 1–11.

Croushore, D., and T. Stark, 2001, "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, 105, 111–130.

Da, Z., 2009, "Cash flow, consumption risk, and the cross-section of stock returns," *Journal of Finance*, 64(2), 923–956.

Da, Z., and M. C. Warachka, 2009, "Cashflow risk, systematic earnings revisions, and the cross-section of stock returns," *Journal of Financial Economics*, 94(3), 448–468.

Drechsler, I., and A. Yaron, 2011, "What's vol got to do with it," *Review of Financial Studies*, 24(1), 1–45.

Epstein, L. G., and S. Zin, 1989, "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Fama, E. F., and K. R. French, 1992, " The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— , 1993, "Common Risk Factors in the Returns on Bonds and Stocks," *Journal of Financial Economics*, 33, 3–56.

——— , 1995, "Size and book-to-market factors in earnings and returns," *Journal of Finance*, 50, 131–155.

——— , 1996, "Multifactor Explanations of Asset Pricing Anomalies," *The Journal of Finance*, 51(1), 55–84.

Ferson, W., S. Nallareddy, and B. Xie, 2012, "The Out-of-Sample Performance of Long-Run Risk Models," *Journal of Financial Economics*, forthcoming.

Gabaix, X., 2008, "Variable rare disasters: A tractable theory of ten puzzles in macro-finance," *The American Economic Review*, 98(2), 64–67.

Gibbons, M. R., and W. Ferson, 1985, "Testing asset pricing models with changing expectations and an unobservable market portfolio," *Journal of Financial Economics*, 14(2), 217 – 236.

Gourio, F., 2008, "Disasters and Recoveries," *American Economic Review, Papers and Proceedings*, 98, 68–73.

Hansen, L., J. Heaton, and N. Li, 2008, "Consumption Strikes Back? Measuring Long Run Risk," *Journal of Political Economy*, 116, 260–302.

He, Z., and A. Krishnamurthy, 2008, "Intermediary asset pricing," working paper, National Bureau of Economic Research.

Jagannathan, R., S. Marakani, H. Takehara, and Y. Wang, 2012, "Calendar Cycles, Infrequent Decisions and the Cross-Section of Stock Returns," *Management Science*, 58(3), 507–522.

Jagannathan, R., and Y. Wang, 2007, "Lazy investors, discretionary consumption, and the cross-section of stock returns," *Journal of Finance*, 62(4), 1623–1661.

Kaltenbrunner, G., and L. Lochstoer, 2010, "Long-run risk through consumption smoothing," *Review of Financial Studies*, 23(8), 3190.

Kan, R., and C. Robotti, 2008, "Specification tests of asset pricing models using excess returns," *Journal of Empirical Finance*, 15(5), 816–838.

Kan, R., and C. Zhang, 1999, "Two-pass tests of asset pricing models with useless factors," *Journal of Finance*, 54, 203–235.

Kapoor, M., and S. Ravi, 2010, "Elasticity of Intertemporal Substitution in Consumption (σ): Empirical Evidence from a Natural Experiment," Working paper.

Kiku, D., 2006, "Is the Value Premium a Puzzle?," USC FBE Seminar Series Paper.

Kleibergen, F., 2009, "Tests of Risk Premia in Linear Factor Models," *Journal of Econometrics*, 149, 149–173.

———— , 2010, "Statistical anomalies in "support" of factor pricing," Working Paper.

Lehmann, B. N., and D. M. Modest, 1988, "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, 21, 213–254.

——— , 2005, "Diversification and the Optimal Construction of Basis Portfolios," *Management Science*, 51, 581–598.

Lettau, M., and J. A. Wachter, 2007, "Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium," *Journal of Finance*, 62, 55–92.

Lewellen, J. W., S. Nagel, and J. A. Shanken, 2006, "A Skeptical Appraisal of Asset Pricing Tests," *SSRN eLibrary*.

Maio, P., and P. Santa-Clara, 2012, "Multifactor models and their consistency with the ICAPM," *Journal of Financial Economics*.

Malloy, C., T. Moskowitz, and A. Vissing-Jørgensen, 2009, "Long-Run Stockholder Consumption Risk and Asset Returns," *Journal of Finance*, 64, 2427–2479.

Marakani, S., 2009, "Long run consumption risks : are they there?," Working Paper.

Mehra, R., and E. Prescott, 1985, "The Equity Premium : A Puzzle," *Journal of Monetary Economics*, 15, 145–61.

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

Newey, W. K., and K. D. West, 1994, "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Parker, J., and C. Julliard, 2005, "Consumption Risk and Cross-Sectional Returns," *Journal of Political Economy*, 113, 185–222.

Parker, J. A., and A. Vissing-Jorgensen, 2010, "The increase in income cyclicality of high-income households and its relation to the rise in top income shares," working paper, National Bureau of Economic Research.

Rietz, T., 1988, "The Equity Premium : a Solution," *Journal of Monetary Economics*, 22, 117–131.

Santos, T., and P. Veronesi, 2009, "Habit Formation, the Cross-Section of Asset Returns and the Cash-Flow Risk Puzzle," Working Paper.

Shanken, J., 1992, "On the estimation of beta-pricing models," *Review of Financial Studies*, 5, 1–33.

Shanken, J., and G. Zhou, 2007, "Estimating and testing beta pricing models: Alternative methods and their performance in simulations," *Journal of Financial Economics*, 84(1), 40–86.

Van Binsbergen, J., and R. Koijen, 2010, "Predictive regressions: A present-value approach," working paper, National Bureau of Economic Research.

Vissing-Jørgensen, A., and O. Attanasio, 2003, "Stock-market participation, intertemporal substitution, and risk-aversion," *American Economic Review*, pp. 383–391.

Weil, P., 1989, "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24(3), 401–421.

——— , 1990, "Non-expected utility in macroeconomics," *Quarterly Journal of Economics*, 105, 29–42.

Whelan, K., 2000, "A Guide to the Use of Chain Aggregated NIPA Data," Note.

Yang, W., 2011, "Long-run risk in durable consumption," *Journal of Financial Economics*, 102, 45–61.

Zhang, L., 2005, "The Value Premium," *Journal of Finance*, 60(1), 67–103.

Zhou, G., and Y. Zhu, 2009, "A Long-run Risks Model with Long- and Short-Run Volatilities: Explaining Predictability and Volatility Risk Premium," *SSRN eLibrary*.

A   Log P/D Ratios in the General Long Run Risk Model

The methodology here closely follows that of Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007) as there are only two common cases where solutions to models with Epstein-Zin-Weil preferences are available in the literature. The first case, which we are interested in here, is when the returns are loglinear in the state variables and the second is when $\psi = 1$.

Let c, $X_i, 1 \leq i \leq n$ and $V_j, 1 \leq j \leq m$ be the log consumption process, n processes that determine it's conditional growth rate and m processes that determine it's conditional growth rate volatility respectively. Let $d_l, l \leq 1 \leq L$ be the log dividend processes of L assets (in general, the lower case variables correspond to the logarithm of the upper case variables). We assume that these quantities follow the processes

$$
\begin{aligned}
c_{t+\Delta t} =& c_t + \left(\mu + \sum_{i=1}^{n} X_{i,t} \right) \Delta t + \sqrt{\sum_{j=1}^{m} \delta_{c,j}^2 V_{j,t}} \left(W_{t+\Delta t} - W_t \right) \\
& - \sum_{k=1}^{m} \varphi_{w,k} \sigma_k \sqrt{V_{k,t}} \left(Z_{k,t+\Delta t} - Z_{k,t} \right)
\end{aligned}
\tag{37}
$$

$$
X_{i,t+\Delta t} = X_{i,t}(1 - \alpha_i \Delta t) + \varphi_{x,i} \sqrt{\sum_{j=1}^{m} \delta_{x,i,j}^2 V_{j,t}} \left(Y_{i,t+\Delta t} - Y_{i,t} \right), 1 \leq i \leq n
\tag{38}
$$

$$
V_{i,t+\Delta t} = V_{i,t} - \kappa_i(V_{i,t} - \bar{V}_i)\Delta t + \sigma_i \sqrt{V_{i,t}} \left(Z_{i,t+\Delta t} - Z_{i,t} \right), 1 \leq i \leq m
\tag{39}
$$

and

$$
\begin{aligned}
d_{l,t+\Delta t} = d_{l,t} &+ \left(\mu_l + \sum_{i=1}^{n} \phi_{l,i} X_{i,t} \right) \Delta t + \pi_{l,c} \left(\Delta c_{t+\Delta t} - \left(\mu + \sum_{i=1}^{n} X_{i,t} \right) \Delta t \right) \\
&+ \sum_{i=1}^{n} \pi_{i,l,x} (X_{i,t+\Delta t} - X_{i,t}(1 - \alpha_i \Delta t)) \\
&+ \sum_{j=1}^{m} \pi_{j,l,w} \sigma_j \sqrt{V_{j,t}} \left(Z_{j,t+\Delta t} - Z_{j,t} \right) \\
&+ \sqrt{\sum_{k=1}^{m} \delta_{l,d,k}^2 V_{k,t}} \, \sigma_{l,d} (B_{l,t+\Delta t} - B_{l,t})
\end{aligned}
\tag{40}
$$

where W, $Y_i, 1 \leq i \leq n$, $Z_j, 1 \leq j \leq m$ are one-dimensional independent Wiener processes, $B_l, 1 \leq l \leq L$ are L Brownian processes which are independent of W, $Y_i, 1 \leq i \leq n$ and $Z_j, 1 \leq j \leq m$ but which can be correlated with each other, and $\sum_{i=1}^{m} \delta_{c,i}^2 = \sum_{j=1}^{m} \delta_{x,i,j}^2 = \sum_{k=1}^{m} \delta_{l,d,k}^2 = 1$. We have written the equations in this form (with the time step being Δt rather than 1) to make the time scale dependence of the parameters explicit so that the connection with the continuous time solution can be made in a straightforward manner. We also define the consumption and dividend variables as rates since they are flow variables. This means, for example, that consumption from time t to $t + \Delta t$ is given by $C_{t+\Delta t} \Delta t$.

Since the consumer preferences are of the Epstein-Zin-Weil type (Epstein and Zin 1989) (Weil 1990)

$$
U_t = ((1 - \delta)(C_t \Delta t)^{\frac{1-\gamma}{\theta}} + \delta E_t [U_{t+\Delta t}^{1-\gamma}]^{\frac{1}{\theta}})^{\frac{\theta}{1-\gamma}}
\tag{41}
$$

where

$$
\theta = \frac{1 - \gamma}{1 - 1/\psi}
\tag{42}
$$

the log stochastic discount factor in discrete time can be written as

$$
m_{t+\Delta t} = \theta \Delta t \log \delta - \frac{\theta}{\psi} \Delta c_{t+\Delta t} + (\theta - 1) r_{c,t+\Delta t}
\tag{43}
$$

A2

where $r_{c,t+\Delta t}$ is the continuously compounded rate of return on the wealth W (which is the asset that delivers a dividend of per capita consumption at every time period) from t to $t+\Delta t$. Since we assume complete markets,

$$E_t[\exp(m_{t+\Delta t} + r_{c,t+\Delta t})] = 1 \tag{44}$$

must hold.

The loglinear approximation pioneered by Campbell and Shiller (1988) allows us to write

$$r_{c,t+\Delta t} = \nu_0 + \nu_1(w_{t+\Delta t} - c_{t+\Delta t}) - (w_t - c_t) + \Delta c_{t+\Delta t} \tag{45}$$

where

$$\nu_0 = \log(\Delta t + \exp(\overline{w-c})) - \nu_1(\overline{w-c}) \approx \exp(\overline{c-w})(1 + (\overline{c-w}))\Delta t \tag{46}$$

$$\nu_1 = \frac{1}{1 + \exp(\overline{c-w})\Delta t} \approx 1 - \exp(\overline{c-w})\Delta t \tag{47}$$

(the approximation holds when Δt is small) where the bar stands for the mean value. We further assume that the log wealth to consumption ratio can be written as

$$w_t - c_t = A_0 + \sum_{i=1}^{n} A_{1,i}X_{i,t} + \sum_{j=1}^{m} A_{2,j}V_{j,t} \tag{48}$$

and justify this below. (This approach is standard and followed by Bansal and Yaron (2004), Bansal, Yaron, and Kiku (2007) and Zhou and Zhu (2009) as the only non-trivial models with Epstein-Zin-Weil preferences which can be solved are those where the consumption to wealth ratio is loglinear in the state variables as above or where $\psi = 1$, as in the model of Hansen, Heaton, and Li (2008)).

Substituting (43), (45) and (48) into (44), using the fact that

$$\log E_t[\exp A(W_{t+\Delta t} - W_t)] = \frac{A^2 \Delta t}{2} \tag{49}$$

for any $A \in \mathbb{R}$ and Wiener process W, and that (44) should hold for any possible attainable combination of state variables (X_i, V_j), we obtain a set of equations which enable us to solve for A_0, $A_{1,i}, 1 \leq i \leq n$ and $A_{2,j}, 1 \leq j \leq m$. The fact that such a set of equations with non-vacuous solutions exist justifies the assumption (48).

The set of equations for $A_{1,i}$ are

$$(1 - \gamma)\Delta t + \theta A_{1,i}(\nu_1(1 - \alpha_i \Delta t) - 1) = 0 \tag{50}$$

so that

$$A_{1,i} = \frac{(1 - \frac{1}{\psi})\Delta t}{1 - \nu_1(1 - \alpha_i \Delta t)} \tag{51}$$

which, in the limit $\Delta t \to 0$, becomes $A_{1,i} = \frac{1 - 1/\psi}{\exp(\overline{c - w}) + \alpha_i}$. This is the same result as that obtained by Zhou and Zhu (2009), where there is only one X variable, once we relate his notation of g_1 for $\exp(\overline{c - w})$ and allow for the negative sign which arises from his definition of A_1 in terms of the consumption to wealth ratio. Once we set $\Delta t = 1$ and relabel ν_1 as κ_1 and α_i as $1 - \rho$ (again, there being only one X state variable) to match the notation of Bansal and Yaron (2004), we find that our result also matches their's.

The analogous set of equations which enables us to solve for $A_{2,j}, 1 \leq j \leq m$ is

$$\begin{aligned}
&\frac{(1 - \gamma)^2 \delta_{c,j}^2 \Delta t}{2} + \theta A_{2,j}(\nu_1(1 - \kappa_j \Delta t) - 1) \\
&+ \frac{\Delta t}{2} \left(\left(\theta \nu_1 \sum_{i=1}^{n} A_{1,i} \varphi_{x,i} \delta_{x,i,j} \right)^2 + (\theta \nu_1 A_{2,j} \sigma_j - (1 - \gamma)\varphi_{w,j}\sigma_j)^2 \right) = 0
\end{aligned} \tag{52}$$

Since these equations are quadratic, there are two solutions for each $A_{2,j}$. However, one of them diverges when $\sigma_j \to 0$. Hence, the other solution is the one which is relevant to the model. The final equation, which allows us to solve for A_0, is

$$\theta \left(\log \delta + \nu_0 + (\nu_1 - 1)A_0 + \nu_1 \sum_{j=1}^m A_{2,j} \kappa_j \Delta t \bar{V}_j \right) + (1 - \gamma)\mu \Delta t = 0 \tag{53}$$

Putting the values for A_0, $A_{1,i}, 1 \leq i \leq n$ and $A_{2,j}, 1 \leq j \leq m$ into (48) and using (45) and (43), we obtain the log stochastic discount factor

$$
\begin{aligned}
m_{t+\Delta t} =& \Delta t \left(\Gamma_0 + \sum_{i=1}^n \Gamma_{1,i} X_{i,t} + \sum_{j=1}^m \Gamma_{2,j} V_{j,t} \right) \\
& - \alpha_c \sqrt{\sum_{j=1}^m \delta_{c,j}^2 V_{j,t}} (W_{t+\Delta t} - W_t) \\
& - \sum_{i=1}^n \alpha_{x,i} \varphi_{x,i} \sqrt{\sum_{j=1}^m \delta_{x,i,j}^2 V_{j,t}} (Y_{i,t+\Delta t} - Y_{i,t}) \\
& - \sum_{j=1}^m \alpha_{v,j} \sigma_j \sqrt{V_{j,t}} (Z_{j,t+\Delta t} - Z_{j,t})
\end{aligned}
\tag{54}
$$

where $\Gamma_0 = \theta \log \delta - \gamma \mu + (\theta - 1)(\frac{\nu_0}{\Delta t} + \frac{\nu_1 - 1}{\Delta t} A_0 + \sum_{j=1}^m \nu_1 \kappa_j A_{2,j} \bar{V}_j)$, $\Gamma_{1,i} = -1/\psi, \forall 1 \leq i \leq n$, $\Gamma_{2,j} = (\theta - 1)A_{2,j}\left(\frac{\nu_1 - 1}{\Delta t} - \kappa_j\right)$, $\alpha_c = \gamma$ and $\alpha_{x,i} = (\theta - 1)\nu_1 A_{1,i} = \frac{\gamma - 1/\psi}{1 - \nu_1(1 - \alpha_i \Delta t)}, \forall 1 \leq i \leq n$. The expression for $\alpha_{v,j}$ is complicated and does not directly concern us here as it's exact form is unimportant for the beta pricing relation.

Using $r_{f,t} = -E_t[m_{t+\Delta t}] - \frac{1}{2}\text{Var}_t[m_{t+\Delta t}]$, we obtain that

$$r_{f,t} = A_{0,f} + \sum_{i=1}^n A_{1,f,i} X_{i,t} + \sum_{j=1}^m A_{2,f,j} V_{j,t} \tag{55}$$

where $A_{0,f} = -\Gamma_0$, $A_{1,f,i} = 1/\psi, \forall 1 \leq i \leq n$ and $A_{2,f,j} = -\Gamma_2 - \frac{1}{2}\left(\alpha_c^2 \delta_{c,j}^2 + \sum_{i=1}^n \alpha_{x,i}^2 \delta_{x,i,j}^2 \varphi_{x,i}^2 + \alpha_{v,j}^2 \sigma_j^2\right), \forall 1 \leq j \leq m$.

A5

Using the process for dividend growth (40), we can use a similar loglinear approximation to write the return for asset l as

$$r_{l,t+\Delta t} = \nu_{0,l} + \nu_{1,l}(p_{l,t+\Delta t} - d_{l,t+\Delta t}) - (p_{l,t} - d_{l,t}) + \Delta d_{l,t+\Delta t} \tag{56}$$

where

$$\nu_{0,l} = \log(\Delta t + \exp(\overline{d_l - p_l})) - \nu_{1,l}(\overline{p_l - d_l})$$
$$\approx \exp(\overline{d_l - p_l})(1 + \overline{d_l - p_l})\Delta t \tag{57}$$

$$\nu_{1,l} = \frac{1}{1 + \exp(\overline{d_l - p_l})\Delta t} \approx 1 - \exp(\overline{d_l - p_l})\Delta t \tag{58}$$

As before, we assume that $\log\left(\frac{P_t}{D_t}\right)$ can be written as

$$\log\left(\frac{P_{l,t}}{D_{l,t}}\right) = p_{l,t} - d_{l,t} = A_{0,l} + \sum_{i=1}^{n} A_{1,l,i} X_{i,t} + \sum_{j=1}^{m} A_{2,l,j} V_{j,t} \tag{59}$$

We put (59) into (56) and use the fact that (44) must hold for any possible attainable combination of state variables (X_i, V_j) to obtain a set of equations which enables us to solve for $A_{0,l}$, $A_{1,l,i}, 1 \leq i \leq n$ and $A_{2,l,j}, 1 \leq j \leq m$. The fact that such a set of equations with non-vacuous solutions exist justifies the assumption (59).

The equations for $A_{1,l,i}, 1 \leq i \leq n, 1 \leq l \leq L$ are

$$(\phi_{l,i} - 1/\psi)\Delta t - A_{1,l,i}(1 - \nu_{1,l}(1 - \alpha_i \Delta t)) = 0 \tag{60}$$

which give

$$A_{1,l,i} = \frac{(\phi_{l,i} - 1/\psi)\Delta t}{1 - \nu_{1,l}(1 - \alpha_i \Delta t)} \tag{61}$$

A6

As with the solution for $A_{1,i}, 1 \leq i \leq n$, this solution agrees with the continuous time one (with $n = 1, m = 2$) of Zhou and Zhu (2009) and the discrete time one (with $n = m = 1$) of Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007).

The equations for $A_{2,l,j}, 1 \leq j \leq m, 1 \leq l \leq L$ are quadratic in nature and fairly complex (as for $A_{2,j}$, the solutions which do not diverge as $\sigma_j \to 0$ are chosen). Since their precise structure is not important for the beta pricing relation, we do not include them for brevity. Similarly, we do not include the equation for $A_{0,l}, 1 \leq l \leq L$.[29]

B   Beta Pricing Relations

B.1   First beta pricing relation

The beta pricing relation can be derived from the fundamental equations of the model (37), (38), (39) and (40) together with the form of the preferences (41). To simplify the notation, we set Δt to 1 in this derivation.

---

[29]They are available upon request from the authors. Also, it must be noted that, as the equations for $A_{2,j}, 1 \leq j \leq m$ and $A_{2,l,j}, 1 \leq j \leq m, 1 \leq l \leq L$ are quadratic in nature, real solutions are not guaranteed. Our numerical experiments indicate that this is not a serious concern as several sets of reasonable parameter values do not give rise to this problem (this is also shown by (Zhou and Zhu 2009)). If this is a concern, we can replace the volatility processes by Ornstein-Uhlenbeck ones as done by (Bansal and Yaron 2004) and (Bansal, Yaron, and Kiku 2007). However, such volatility processes suffer from the problem of admitting negative values even in continuous time. This can be quite serious, even for some common parameter values, as pointed out by (Beeler and Campbell 2012). The square root processes used here can also give rise to negative values in discrete time but the probability of this occurring for reasonable parameter values is minuscule and our numerical experiments confirm this. Since both ways of modeling volatility have issues but have received wide attention in the literature and there is no known alternative for which analytical solutions can be derived, we use results which hold for both of them.

Theorem 3. *The model defined by (37), (38), (39) and (40) implies the beta pricing relation*

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t} - \sum_{j=1}^{m}\beta_{l,V_j}V_{j,t}\right] =$$

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^2\varphi_{x,i}^2 + \alpha_{v,j}\beta_{l,Z_j}\sigma_j^2 + \alpha_c\beta_{l,c}\delta_{c,j}^2\right)\bar{V}_j \qquad (62)$$

$$= \sum_{i=1}^{n}\beta_{l,Y_i}\lambda_{Y_i} + \sum_{j=1}^{m}\beta_{l,Z_j}\lambda_{Z_j} + \beta_{l,c}\lambda_c$$

where

$$\lambda_{Y_i} = \alpha_{x,i}\varphi_{x,i}^2\sum_{j=1}^{m}\delta_{x,i,j}^2\bar{V}_j \qquad (63)$$

$$\lambda_{Z_j} = \alpha_{v,j}\sigma_j^2\bar{V}_j \qquad (64)$$

$$\lambda_c = \alpha_c\sum_{j=1}^{m}\delta_{c,j}^2\bar{V}_j \qquad (65)$$

where the $\beta_{l,V_j}, 1 \leq j \leq m$, $\beta_{l,Y_i}, 1 \leq i \leq n$, $\beta_{l,Z_j}, 1 \leq j \leq m$ and $\beta_{l,c}$ are, respectively, the multiple regression coefficients obtained on regressing $r_{l,t+1} - r_{f,t}$ on $V_{j,t}, 1 \leq j \leq m$, $Y_{i,t+1}, 1 \leq 1 \leq n$, $Z_{j,t+1}, 1 \leq j \leq m$ and the residuals obtained on regressing Δc_{t+1} against $X_{i,t}, 1 \leq i \leq n$.

This can be written in terms of the $n+m$ log price dividend ratio factors $F_i, 1 \leq i \leq m+n$ and their innovations $IF_i, 1 \leq n+m$ as

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t} - \sum_{i=1}^{m+n}\beta_{l,F_i}F_{i,t}\right] =$$

$$\sum_{i=1}^{m+n}\beta_{l,IF_i}\lambda_{IF_i} + \beta_{l,c}\lambda_c \qquad (66)$$

A8

When there is only one volatility factor, the beta pricing relation can also be written as

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t}\right]$$

$$= \beta_{l,V}^2 \text{Var}[V_t] + \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i}^2\varphi_{x,i}^2 + \alpha_v\beta_{l,Z}\sigma^2 + \alpha_c\beta_{l,c}\delta_c^2\right)\bar{V} \quad (67)$$

$$= \beta_{l,V}^2\lambda_V + \sum_{i=1}^{n}\beta_{l,Y_i}\lambda_{Y_i} + \beta_{l,Z}\lambda_Z + \beta_{l,c}\lambda_c$$

with the βs defined in a similar manner as above and where

$$\lambda_V = \text{Var}[V_t] \quad (68)$$

$$\lambda_{Y_i} = \alpha_{x,i}\varphi_{x,i}^2\bar{V} \quad (69)$$

$$\lambda_Z = \alpha_v\sigma^2\bar{V} \quad (70)$$

$$\lambda_c = \alpha_c\bar{V} \quad (71)$$

This can be written in terms of the $n + m$ log price dividend ratio factors $F_i, 1 \le i \le n+1$ and their innovations $IF_i, 1 \le n+1$ as

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t}\right]$$

$$= \beta_{l,F_k}^2\lambda_{F_k} + \sum_{i=1}^{n+1}\beta_{l,IF_i}\lambda_{IF_i} + \beta_{l,c}\lambda_c \quad (72)$$

where the βs are as defined above and F_k is some log price dividend ratio factor that has non-zero correlation with the volatility factor.

Proof. We use (56) and (59) to write

$$
\begin{aligned}
r_{l,t+1} =&\, \nu_{0,l} + \nu_{1,l}\left(A_{0,l} + \sum_{i=1}^{n} A_{1,l,i}X_{i,t+1} + \sum_{j=1}^{m} A_{2,l,j}V_{j,t+1} \right) \\
&- \left(A_{0,l} + \sum_{i=1}^{n} A_{1,l,i}X_{i,t} + \sum_{j=1}^{m} A_{2,l,j}V_{j,t} \right) + \Delta d_{l,t+1}
\end{aligned}
\tag{73}
$$

Using (37), (38), (39) and (40), (73) becomes

$$
\begin{aligned}
r_{l,t+1} =&\, \nu_{0,l} + (\nu_{1,l} - 1)A_{0,l} + \nu_{1,l}\sum_{j=1}^{m} A_{2,l,j}\kappa_j \bar{V}_j + \mu_l \\
&+ \sum_{i=1}^{n} \left(A_{1,l,i}(\nu_{1,l}(1-\alpha_i) - 1) + \phi_{l,i} \right) X_{i,t} \\
&+ \sum_{j=1}^{m} A_{2,l,j}(\nu_{1,l}(1-\kappa_j) - 1)V_{j,t} \\
&+ \pi_{l,c} \boxed{\left(\sqrt{\sum_{j=1}^{m} \delta_{c,j}^2 V_{j,t}}(W_{t+1} - W_t) - \sum_{j=1}^{m} \varphi_{w,j}\sigma_j\sqrt{V_{j,t}}\left(Z_{j,t+1} - Z_{j,t} \right) \right)}_{consumption\ innovation} \\
&+ \sum_{i=1}^{n}(\pi_{i,l,x} + \nu_{1,l}A_{1,l,i}) \boxed{\varphi_{x,i}\sqrt{\sum_{j=1}^{m} \delta_{x,i,j}^2 V_{j,t}}(Y_{i,t+1} - Y_{i,t})}_{X_i\ innovation} \\
&+ \nu_{1,l}\sum_{j=1}^{m} A_{2,l,j}\boxed{\sigma_j\sqrt{V_{j,t}}(Z_{j,t+1} - Z_{j,t})}_{V_j\ innovation} \\
&+ \boxed{\sigma_{l,d}\sqrt{\sum_{j=1}^{m} \delta_{l,d,j}^2 V_{j,t}}(B_{l,t+\Delta t} - B_{l,t})}_{residual}
\end{aligned}
\tag{74}
$$

A10

which can be written as

$$
\begin{aligned}
r_{l,t+1} =& \beta_{l,0} + \sum_{i=1}^{n} \beta_{l,X_i} X_{i,t} + \sum_{j=1}^{m} \beta_{l,V_j} V_{j,t} + \beta_{l,c}(\Delta c_{t+1} - E_t[\Delta c_{t+1}]) \\
&+ \sum_{i=1}^{n} \beta_{l,Y_i}(X_{i,t+1} - E_t[X_{i,t+1}]) + \sum_{j=1}^{m} \beta_{l,Z_j}(V_{j,t+1} - E_t[V_{j,t+1}]) + \epsilon_{l,t+1}
\end{aligned}
\tag{75}
$$

where

$$
\beta_{l,0} = \nu_{0,l} + (\nu_{1,l} - 1)A_{0,l} + \nu_{1,l} \sum_{j=1}^{m} A_{2,l,j}\kappa_j \bar{V}_j + \mu_l
\tag{76}
$$

$$
\beta_{l,X_i} = A_{1,l,i}(\nu_{1,l}(1 - \alpha_i) - 1) + \phi_{l,i}
\tag{77}
$$

$$
\beta_{l,V_j} = A_{2,l,j}(\nu_{1,l}(1 - \kappa_j) - 1)
\tag{78}
$$

$$
\beta_{l,c} = \pi_{l,c}
\tag{79}
$$

$$
\beta_{l,Y_i} = \pi_{i,l,x} + \nu_{1,l}A_{1,l,i}
\tag{80}
$$

$$
\beta_{l,Z_j} = \nu_{1,l}A_{2,l,j}
\tag{81}
$$

$$
\epsilon_{l,t+1} = \sigma_{l,d}\sqrt{\sum_{j=1}^{m} \delta_{l,d,j}^2 V_{j,t}}(B_{l,t+\Delta t} - B_{l,t})
\tag{82}
$$

The standard asset pricing relation

$$
E_t[\exp(m_{t+1} + r_{l,t+1})] = 1
\tag{83}
$$

together with conditional normality implies that

$$
E_t[m_{t+1} + r_{l,t+1}] + \frac{1}{2}\text{Var}_t[m_{t+1} + r_{l,t+1}] = 0
\tag{84}
$$

A11

From (75) and (43), we get

$$E_t[m_{t+1} + r_{l,t+1}] = (\Gamma_0 + \beta_{l,0}) + \sum_{i=1}^{n}(\Gamma_{1,i} + \beta_{l,X_i})X_{i,t} + \sum_{j=1}^{m}(\Gamma_{2,j} + \beta_{l,V_j})V_{j,t} \tag{85}$$

and

$$\text{Var}_t[m_{t+1} + r_{l,t+1}] = \sum_{j=1}^{m}\left(\sum_{i=1}^{n}(\beta_{l,Y_i} - \alpha_{x,i})^2\delta_{x,i,j}^2\varphi_{x,i}^2 + (\beta_{l,Z_j} - \alpha_{v,j})^2\sigma_j^2 + (\beta_{l,c} - \alpha_c)^2\delta_{c,j}^2\right)V_{j,t} \tag{86}$$

Putting (84), (85) and (86) together and using the fact that the asset pricing relation must hold for any values of $X_{i,t}$ and $V_{j,t}$, we get the following relations

$$\beta_{l,0} = -\Gamma_0 = A_{0,f} \tag{87}$$

$$\beta_{l,X_i} = -\Gamma_{1,i} = 1/\psi = A_{1,f} \tag{88}$$

$$\beta_{l,V_j} = -\Gamma_{2,j} - \frac{1}{2}\left(\sum_{i=1}^{n}(\beta_{l,Y_i} - \alpha_{x,i})^2\delta_{x,i,j}^2\varphi_{x,i}^2 + (\beta_{l,Z_j} - \alpha_{v,j})^2\sigma_j^2 + (\beta_{l,c} - \alpha_c)^2\delta_{c,j}^2\right) \tag{89}$$

$$= A_{2,f} + \left(\sum_{i=1}^{n}\alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^2\varphi_{x,i}^2 + \alpha_{v,j}\beta_{l,Z_j}\sigma_j^2 + \alpha_c\beta_{l,c}\delta_{c,j}^2\right) - \frac{1}{2}\left(\beta_{l,Y_i}^2\delta_{x,i,j}^2\varphi_{x,i}^2 + \beta_{l,Z_j}^2\sigma_j^2 + \beta_{l,c}^2\delta_{c,j}^2\right)$$

Hence,

$$r_{l,t+1} - r_{f,t} = (\beta_{l,0} - A_{0,f}) + \sum_{i=1}^{n}(\beta_{l,X_i} - A_{1,f,i})X_{i,t} + \sum_{j=1}^{m}(\beta_{l,V_j} - A_{2,f,j})V_{j,t}$$

$$+ \beta_{l,c}(\Delta c_{t+1} - E_t[\Delta c_{t+1}]) + \sum_{i=1}^{n}\beta_{l,Y_i}(X_{i,t+1} - E_t[X_{i,t+1}]) \tag{90}$$

$$+ \sum_{j=1}^{m}\beta_{l,Z_j}(V_{j,t+1} - E_t[V_{j,t+1}]) + \epsilon_{l,t+1}$$

A12

or

$$r_{l,t+1} - r_{f,t} = \sum_{j=1}^{m} \beta_{l,V_j}^{e} V_{j,t} + \beta_{l,c}(\Delta c_{t+1} - E_t[\Delta c_{t+1}]) + \sum_{i=1}^{n} \beta_{l,Y_i}(X_{i,t+1} - E_t[X_{i,t+1}])$$
$$+ \sum_{j=1}^{m} \beta_{l,Z_j}(V_{j,t+1} - E_t[V_{j,t+1}]) + \epsilon_{l,t+1}$$

(91)

where

$$\beta_{l,V_j}^{e} = \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^{2}\varphi_{x,i}^{2} + \alpha_{v,j}\beta_{l,Z_j}\sigma_{j}^{2} + \alpha_{c}\beta_{l,c}\delta_{c,j}^{2} \right)$$
$$- \frac{1}{2}\left(\beta_{l,Y_i}^{2}\delta_{x,i,j}^{2}\varphi_{x,i}^{2} + \beta_{l,Z_j}^{2}\sigma_{j}^{2} + \beta_{l,c}^{2}\delta_{c,j}^{2} \right)$$

(92)

$$\implies \sum_{j=1}^{m} \beta_{l,V_j}^{e} V_{j,t} = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^{2}\varphi_{x,i}^{2} + \alpha_{v,j}\beta_{l,Z_j}\sigma_{j}^{2} + \alpha_{c}\beta_{l,c}\delta_{c,j}^{2} \right) V_{j,t}$$
$$- \frac{1}{2}\text{Var}_t[r_{l,t+1} - r_{f,t}]$$

(93)

Combining (93) with (91) gives

$$E_t[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}_t[r_{l,t+1} - r_{f,t}] = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^{2}\varphi_{x,i}^{2} + \alpha_{v,j}\beta_{l,Z_j}\sigma_{j}^{2} + \alpha_{c}\beta_{l,c}\delta_{c,j}^{2} \right) V_{j,t}$$

(94)

Taking expectations gives

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}E\left[\text{Var}_t[r_{l,t+1} - r_{f,t}]\right] =$$
$$\sum_{j=1}^{m} \left(\sum_{i=1}^{n} \alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^{2}\varphi_{x,i}^{2} + \alpha_{v,j}\beta_{l,Z_j}\sigma_{j}^{2} + \alpha_{c}\beta_{l,c}\delta_{c,j}^{2} \right) \bar{V}_j$$

(95)

or, equivalently

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t} - \sum_{j=1}^{m}\beta_{l,V_j}V_{j,t}\right] =$$

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\alpha_{x,i}\beta_{l,Y_i}\delta_{x,i,j}^2\varphi_{x,i}^2 + \alpha_{v,j}\beta_{l,Z_j}\sigma_j^2 + \alpha_c\beta_{l,c}\delta_{c,j}^2\right)\bar{V}_j \qquad (96)$$

$$= \sum_{i=1}^{n}\beta_{l,Y_i}\lambda_{Y_i} + \sum_{j=1}^{m}\beta_{l,Z_j}\lambda_{Z_j} + \beta_{l,c}\lambda_c$$

where

$$\lambda_{Y_i} = \alpha_{x,i}\varphi_{x,i}^2\sum_{j=1}^{m}\delta_{x,i,j}^2\bar{V}_j \qquad (97)$$

$$\lambda_{Z_j} = \alpha_{v,j}\sigma_j^2\bar{V}_j \qquad (98)$$

$$\lambda_c = \alpha_c\sum_{j=1}^{m}\delta_{c,j}^2\bar{V}_j \qquad (99)$$

Since the price dividend ratios are affinely related to the X and V state variables by (59), this is equivalent to a beta pricing relation in the innovations to the log price-dividend ratio factors. The computation of the second expression in the left hand side of (96) can be done by noting that the term in brackets is the residual of regressing $r_{l,t+1} - r_{f,t}$ against V_t and is equal to the residual of regressing $r_{l,t+1} - r_{f,t}$ against the lagged log price dividend ratio factors.

An alternative approach using betas instead of residual returns can be developed when there is one volatility factor. We now proceed to do so.

When there is only one volatility factor (i.e. $m = 1$), this simplifies to

$$E[r_{l,t+1} - r_{f,t}] + \frac{1}{2}\text{Var}\left[r_{l,t+1} - r_{f,t}\right] =$$

$$\beta_{l,V}^2\text{Var}[V_t] + \left(\sum_{i=1}^{n}\alpha_{x,i}\beta_{l,Y_i}\delta_{x,i}^2\varphi_{x,i}^2 + \alpha_v\beta_{l,Z}\sigma^2 + \alpha_c\beta_{l,c}\delta_c^2\right)\bar{V} \qquad (100)$$

A14

after dropping the superfluous j subscript. Since $E[r_{l,t+1} - r_{f,t}] = \beta_{l,V}V_t$ from (91), a scale multiple of $\beta_{l,V}$ can be obtained by regressing $r_{l,t+1} - r_{f,t}$ against a lagged log price divided ratio factor which is known to be correlated with the volatility factor.[30] □

It can be shown that the above beta pricing relation also holds when an Ornstein-Uhlenbeck process is used for the volatilities instead of Bessel processes as in (Bansal and Yaron 2004), (Constantinides and Ghosh 2008) and (Bansal, Yaron, and Kiku 2007).[31]

B.2 Second beta pricing relation

Theorem 4. *The model defined by (37), (38), (39) and (40) implies the beta pricing relation*

$$E[r_{l,t+1} - r_{f,t}] = \sum_{j=1}^{m} \beta_{l,V_j}\bar{V}_j \tag{101}$$

where β_{l,V_j} is the coefficient obtained upon regressing $r_{l,t+1} - r_{f,t}$ on $V_{j,t}$.

This can be written in terms of the $n + m$ log price dividend ratio factors $F_i, 1 \le i \le m+n$ as

$$E[r_{l,t+1} - r_{f,t}] = \sum_{i=1}^{m+n} \beta_{l,F_i}\lambda_{F_i} \tag{102}$$

Proof. Taking the expectation at time t of (91), we obtain

$$E_t[r_{l,t+1} - r_{f,t}] = \sum_{j=1}^{m} \beta_{l,V_j}V_{j,t} \tag{103}$$

---

[30]Specifically, we have $\log\left(\frac{P_{l,t}}{D_{l,t}}\right) = p_{l,t} - d_{l,t} = A_{0,l} + \sum_{i=1}^{n} A_{1,l,i}X_{i,t} + A_{2,l}V_t$ which can be inverted to give $V_t = k_0 + \sum_{i=1}^{n+1} k_i \log\left(\frac{P_{i,t}}{D_{i,t}}\right)$ for some $n + 1$ log price to dividend ratios so that $E[r_{l,t+1} - r_{f,t}] = \beta_{l,V}\left(k_0 + \sum_{i=1}^{n+1} k_i \log\left(\frac{P_{i,t}}{D_{i,t}}\right)\right)$.

[31]The proof is available upon request.

A15

Taking the unconditional expectation, we obtain the result

$$E[r_{l,t+1} - r_{f,t}] = \sum_{j=1}^{m} \beta_{l,V_j} \bar{V}_j \tag{104}$$

Since the price dividend ratios are affinely related to the X and V state variables by (59), this is equivalent to a beta pricing relation in the log price-dividend ratio factors which can be explicitly written as

$$E[r_{l,t+1} - r_{f,t}] = \sum_{i=1}^{m+n} \beta_{l,F_i} \lambda_{F_i} \tag{105}$$

□

This beta pricing relation does *not* hold when the volatility processes are Ornstein-Uhlenbeck rather than Bessel. It is, therefore, not an implication of the models of Bansal and Yaron (2004) and Bansal, Yaron, and Kiku (2007).

C Testing Long Run Risk Models : Monte Carlo Evidence

C.1 The Model

For the purpose of analyzing the performance of the asset pricing tests, we use the long run risk model of (Bansal and Yaron 2004). In this model, the per capita consumption and dividend growth rates Δc and Δd (for M assets indexed by l) and their common persistent component x are assumed to follow the processes assumed by (Bansal and Yaron 2004)

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \tag{106}$$

$$x_{t+1} = \rho x_t + \varphi_x \sigma_t e_{t+1} \tag{107}$$

$$\Delta d_{l,t+1} = \mu_{l,d} + \phi_l x_t + \varphi_{l,d} \sigma u_{l,t+1}, 1 \le l \le M \tag{108}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1} \tag{109}$$

where the shocks e_{t+1}, η_{t+1} and w_{t+1} are taken to be independent standard normals for parsimony. $u_{l,t+1}$ is a vector of normally distributed shocks with covariance V_u which is independent of e, η and w. In the simulations, V_u is set so as to fit the factor structure of returns. (Note that we follow the convention that lowercase characters stand for the logarithm of quantities denoted by the corresponding uppercase characters.)

Consumers in the model have Epstein-Zin-Weil preferences (Epstein and Zin 1989) (Weil 1990)

$$U_t = ((1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta E_t[U_{t+1}^{1-\gamma}]^{\frac{1}{\theta}})^{\frac{\theta}{1-\gamma}} \tag{110}$$

with $\gamma > 1/\psi$. This implies that they prefer early resolution of uncertainty and that persistent consumption and volatility shocks have a high market price of risk. This high price of risk results in a high equity premium and low risk-free rate. With these preferences, asset returns satisfy

$$E_t \left[\delta^\theta \left(\frac{C_{t+1}}{C_t} \right)^{-\theta/\psi} R_{a,t+1}^{-(1-\theta)} R_{i,t+1} \right] = 1 \tag{111}$$

where C is per capita consumption, R_a is the gross return on an asset that pays a dividend of per capita consumption, R_i is the asset return, $0 < \delta < 1$ is the time discount factor, γ is the relative risk aversion, ψ is the intertemporal elasticity of substitution (IES) and θ is defined to be

$$\theta = \frac{1-\gamma}{1 - \frac{1}{\psi}} \tag{112}$$

The log P/D ratios of assets in this economy have a factor structure (within the loglinear approximation) with the factors being x_t and σ_t^2. In other words, if $z_{i,t}$ is the log P/D ratio of asset i, we have

$$z_{i,t} = A_{0,i} + A_{1,i}x_t + A_{2,i}\sigma_t^2 \tag{113}$$

This is shown for this particular model by (Bansal and Yaron 2004) and similar results for related models are shown by (Bansal, Yaron, and Kiku 2007), (Drechsler and Yaron 2011),

A17

(Zhou and Zhu 2009), (Ferson, Nallareddy, and Xie 2012) and in A of this paper. Since the dividend processes of the assets are specified in this model, the relation above gives the time series of their prices for a given realization of the random variables. Hence, the prices and other quantities of interest in this economy are readily simulated.

C.2 Monte Carlo Simulation of the Model

The global parameters summarized in table 16 while the asset specific parameters are chosen to replicate the factor structure of log price dividend ratios and returns.[32] We first note that these parameters generate economic moments (calculated from 500 simulations of the long run risk economy) which are roughly in line with the values observed in post-1942 (to account for the structural break identified by Marakani (2009)) US consumption and return data as shown in table 18. When realistic noise is added to the log P/D ratios as described below, they are also compatible with the predictability of real time consumption growth in the data as seen from the numbers in table 19. One moment which does not match well is the standard deviation of the real risk free rate which is *much* smaller in the simulations than in the data. This, however, as argued by Beeler and Campbell (2012), points to a strength rather than a weakness of the long run risk model as most models struggle to make this quantity low enough. Further, as we argue in the next section, this quantity is very noisily measured which means that the reported standard deviation would be significantly larger than the actual one.[33]

The scaled eigenvalues of the covariance matrix of the post-1942 continuously compounded excess returns of the 25 Fama-French portfolios sorted on the basis of size and book to market ratio are tabulated in table 20 together with the mean, 5th and 95th percentiles of the corresponding values obtained in 500 simulations of the above model for the same time period

---

[32]They are available upon request from the authors.

[33]Measurement error (in either inflation or dividends) can also account for the somewhat low standard deviation of real dividend growth of the portfolios in the simulations.

Parameter	Value
μ	0.02
σ	0.012
ρ	0.85
φ_x	0.45
ν	0.99
σ_w	10^{-5}
γ	25
ψ	1.5
δ	0.994

Table 16: Global parameters for the simulation (the time unit is one year). μ represents the unconditional mean of consumption growth, σ it's conditional volatility, ρ the first order autocorrelation of the long run risk state variable x, φ_x the conditional volatility of x in relation to that of consumption growth, ν the first order autocorrelation of volatility, σ_w the volatility of volatility, γ the relative risk aversion, ψ the elasticity of intertemporal substitution and δ the time preference.

(65 years).[34] Since the first few eigenvalues, which are of principal interest, are very similar to those in the data, the model replicates the observed factor structure of excess returns quite well.

The model also replicates the observed factor structure of log P/D ratios fairly well. This is best seen from the normalized eigenvalues for the covariance matrix of the log P/D ratios of the assets, both from the data as well as the simulations, which are tabulated in table 21. The model's two factor structure is highly evident here as all the eigenvalues after the second one are zero. To better reflect the data and investigate the possible consequences of the inclusion of small, irrelevant factors into the long run risk model, we added white noise with a variance of 20% of the simulated values to the log P/D ratios. The introduction of this noise can also be thought of as representing measurement error in the prices or dividends brought about due to liquidity issues or other market imperfections. The normalized eigenvalues after adding this

---

[34]The model was actually simulated for 165 years with the data for the first 100 years being discarded so as to minimize the effect of the assumed initial values of the dynamic quantities.

Parameters for the asset dividend growths

l	$\mu_{l,d}$	ϕ_l	$\varphi_{l,d}$
1	−0.0286	1.7834	19.1677
2	0.0889	3.7689	21.7081
3	0.0160	3.2545	19.4655
4	0.0456	3.4405	23.5766
5	0.0471	2.6758	24.0000
6	0.0907	4.6342	16.6065
7	0.0778	5.8088	16.3543
8	0.0457	2.4918	8.5237
9	0.0928	9.5089	24.0000
10	−0.0145	5.5979	24.0000
11	−0.0012	4.8912	24.0000
12	0.0821	8.5459	22.0032
13	0.0556	10.9271	8.9635
14	0.0272	6.0810	21.8607
15	0.0926	5.1230	24.0000
16	0.0454	5.1540	6.0000
17	0.0327	3.0965	21.1709
18	0.0317	3.3548	16.4485
19	0.0147	3.5232	23.0091
20	0.0619	3.3028	6.6980
21	0.0167	2.5690	12.5081
22	0.0421	10.8271	6.0000
23	0.0901	3.7845	11.6097
24	0.0436	2.5953	24.0000
25	0.0788	3.7323	11.0877

Table 17: Asset-specific parameters for the simulation. The assets are indexed by l. $\mu_{l,d}$ represents the unconditional mean of the dividend growth for asset l, ϕ_l the dependence of predictable dividend growth on the long run risk state variable x and $\varphi_{l,d}$ the idiosyncratic volatility of dividend growth.

noise are summarized in table 22. From it, we see that the model is able to replicate the key elements of this factor structure after adding the noise.[35]

(Albuquerque, Eichenbaum, and Rebelo 2012) have recently pointed out that some versions of the long run risk model imply very high consumption-return correlations over long time scales. We plot this correlation for our simulated model as a function of the time frame in

[35]Note that it is not necessary to replicate the features of the small factors as these represent a very small fraction of the variance and are not economically interesting.

Moment	Data	Simulation mean	5th percentile	95th percentile
$E[\Delta c_t]$	0.0199	0.0200	0.0153	0.0246
$\text{Std}[\Delta c_t]$	0.0136	0.0151	0.0105	0.0194
$AC(1)[\Delta c_t]$	0.243	0.320	0.148	0.488
$E[r_{f,t}]$	0.0059	0.0035	-0.0012	0.0079
$\text{Std}[r_{f,t}]$	0.0343	0.0067	0.0045	0.0089
$\text{Min}[r_{l,t} - r_{f,t}]$	0.010	0.018	-0.012	0.049
$\text{Max}[r_{l,t} - r_{f,t}]$	0.133	0.209	0.131	0.292
$\text{Min}\, E[\Delta d_{l,t}]$	-0.023	-0.030	-0.062	0.002
$\text{Max}\, E[\Delta d_{l,t}]$	0.105	0.104	0.070	0.149
$\text{Min}\, \text{Std}[\Delta d_{l,t}]$	0.087	0.085	0.075	0.095
$\text{Max}\, \text{Std}[\Delta d_{l,t}]$	0.385	0.306	0.279	0.333

Table 18: Moments implied by the model and from the data. The model implied moments are obtained from 500 simulations for a time period equal to that of the data (65 years). The model was actually simulated for 165 years with the first 100 years of data being discarded to ensure that arbitrary initial values did not have any impact on the results.

R^2 obtained on regressing consumption growth against the lagged values of the first two principal components of the log P/D ratios

Data	Simulation mean	5th percentile	95th percentile
17.4%	32.6%	10.6%	55.2%

Table 19: Predictability of consumption growth in the model and in the data. For the data, we use real time consumption growth as a measure of consumption growth. The results for the model are derived from 1000 simulations over 165 years with the data for the first 100 years being dropped so as to limit the impact of initial values on the numbers.

figure 8. We see from it that our simulated model is compatible with a low value for this correlation even over long time frames for the given length of the data (65 years).[36]

We thus see that the long run risk model being simulated here is compatible not only with many of the important observed moments of macroeconomic quantities but also with the observed factor structure of excess returns and P/D ratios. Given this, it is interesting to examine the performance of different asset pricing tests for long run risk models within the context of these simulations. This will enable the study of the effect of finite sample size and measurement noise on the efficacy of these tests and will point to the choice of test to be used

---

[36]Further details are available on request.

Eigenvalues of the covariance matrix of excess returns

Data	Simulation mean	5th percentile	95th percentile
1.00000	1.00000	1.00000	1.00000
0.06052	0.06171	0.04705	0.07889
0.04741	0.03926	0.02994	0.04970
0.01280	0.01135	0.00871	0.01403
0.00823	0.00807	0.00637	0.01010
0.00626	0.00667	0.00531	0.00824
0.00535	0.00573	0.00455	0.00711
0.00389	0.00497	0.00399	0.00613
0.00339	0.00433	0.00351	0.00541
0.00316	0.00375	0.00302	0.00460
0.00288	0.00331	0.00267	0.00403
0.00231	0.00294	0.00240	0.00359
0.00207	0.00263	0.00214	0.00324
0.00200	0.00236	0.00191	0.00289
0.00149	0.00213	0.00170	0.00265
0.00142	0.00191	0.00152	0.00234
0.00132	0.00171	0.00136	0.00212
0.00108	0.00151	0.00118	0.00186
0.00099	0.00132	0.00106	0.00164
0.00097	0.00112	0.00087	0.00139
0.00074	0.00076	0.00059	0.00095
0.00067	0.00058	0.00045	0.00075
0.00056	0.00040	0.00030	0.00050
0.00045	0.00030	0.00023	0.00038
0.00043	0.00023	0.00017	0.00029

Table 20: Eigenvalues of the covariance matrix of the continuously compounded excess returns of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008 as well as those obtained by simulating the model for a time span equal to that of the data (65 years). The model was actually simulated for 165 years with the first 100 years of data being discarded to ensure that arbitrary initial values did not have any impact on the results.

in this paper. Since we are particularly interested in examining the impact of measurement noise on these tests, we first turn to the task of establishing a reasonable estimate for the size of this noise for two important quantities in long run risk models, the consumption growth and the real risk free rate.

Eigenvalues of the covariance matrix of log P/D ratios

Data	Simulation mean	5th percentile	95th percentile
1.00000	1.00000	1.00000	1.00000
0.06041	0.03598	0.01272	0.07067
0.01669	0.00000	0.00000	0.00000
0.01169	0.00000	0.00000	0.00000
0.00627	0.00000	0.00000	0.00000
0.00522	0.00000	0.00000	0.00000
0.00494	0.00000	0.00000	0.00000
0.00318	0.00000	0.00000	0.00000
0.00245	0.00000	0.00000	0.00000
0.00238	0.00000	0.00000	0.00000
0.00215	0.00000	0.00000	0.00000
0.00168	0.00000	0.00000	0.00000
0.00137	0.00000	0.00000	0.00000
0.00101	0.00000	0.00000	0.00000
0.00094	0.00000	0.00000	0.00000
0.00085	0.00000	0.00000	0.00000
0.00072	0.00000	0.00000	0.00000
0.00063	0.00000	0.00000	0.00000
0.00052	0.00000	0.00000	0.00000
0.00049	0.00000	0.00000	0.00000
0.00046	0.00000	0.00000	0.00000
0.00040	0.00000	0.00000	0.00000
0.00028	0.00000	0.00000	0.00000
0.00022	0.00000	0.00000	0.00000
0.00018	0.00000	0.00000	0.00000

Table 21: Eigenvalues of the covariance matrix of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008 as well as those obtained by simulating the model for an equivalent time period (65 years). The model was actually simulated for 165 years with the first 100 years of data being discarded to ensure that arbitrary initial values did not have any impact on the results.

C.3 Measurement Noise

We do so by analyzing the degree of correlation between different measures for the same fundamental macroeconomic quantities. For consumption growth, we use the estimates of consumption growth derived from the continuously revised NIPA tables as well as those from

Eigenvalues of the covariance matrix of noisy log P/D ratios

Data	Simulation mean	5th percentile	95th percentile
1.00000	1.00000	1.00000	1.00000
0.06041	0.04536	0.02144	0.08128
0.01669	0.01451	0.01323	0.01577
0.01169	0.01337	0.01234	0.01442
0.00627	0.01251	0.01160	0.01352
0.00522	0.01179	0.01095	0.01269
0.00494	0.01114	0.01046	0.01183
0.00318	0.01057	0.00989	0.01132
0.00245	0.01003	0.00936	0.01071
0.00238	0.00952	0.00877	0.01022
0.00215	0.00904	0.00842	0.00972
0.00168	0.00859	0.00800	0.00918
0.00137	0.00813	0.00757	0.00872
0.00101	0.00771	0.00712	0.00837
0.00094	0.00730	0.00677	0.00786
0.00085	0.00692	0.00639	0.00751
0.00072	0.00652	0.00603	0.00706
0.00063	0.00615	0.00567	0.00664
0.00052	0.00579	0.00533	0.00627
0.00049	0.00541	0.00496	0.00588
0.00046	0.00506	0.00464	0.00554
0.00040	0.00470	0.00429	0.00514
0.00028	0.00433	0.00388	0.00477
0.00022	0.00394	0.00350	0.00437
0.00018	0.00345	0.00295	0.00390

Table 22: Eigenvalues of the covariance matrix of the log P/D ratios of the 25 Fama-French portfolios formed on the basis of size and book to market ratio as well as those obtained by simulating the model for an equivalent period (65 years) and adding some noise to the result. The model was actually simulated for 165 years with the first 100 years of data being discarded to ensure that arbitrary initial values did not have any impact on the results.

the real time database maintained by the Federal Reserve Bank of St. Louis (described in detail by Croushore and Stark (2001)). Regressing these estimates against each other leads to the results in table 23. The R^2 of 67% or about $\frac{2}{3}$ indicates that the variance of measurement noise in consumption growth is about half of the variance of actual consumption growth. We
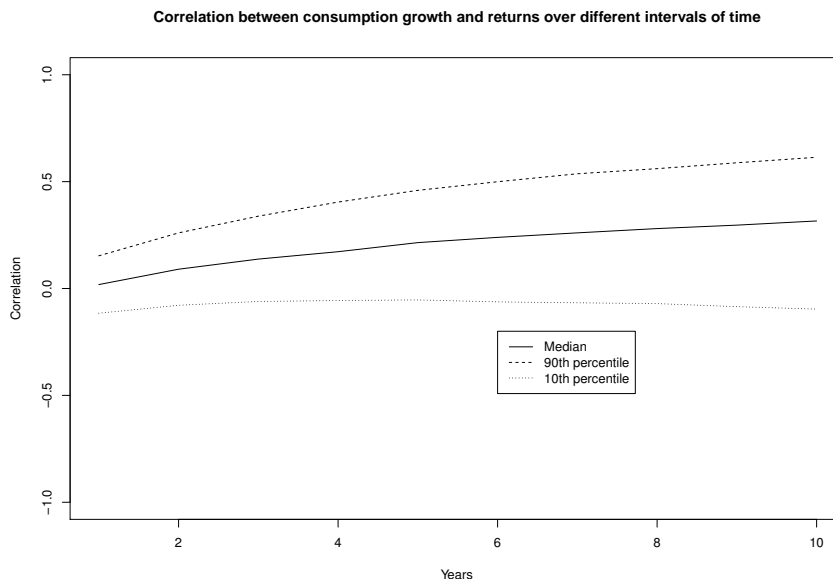
Figure 8: Consumption growth - return correlations as a function of the time frame for the model simulated in this appendix. The simulations are for the same period as the data (65 years). The model was actually simulated for 165 years with the first 100 years of data being discarded to ensure that arbitrary initial values did not have any impact on the results.

thus simulate measured consumption growth as actual consumption growth plus iid noise with half it's realized variance in that simulation.

Regression of Δc_t against Δc_t^{RT}

Coefficient	Estimate (Std. Err.)
Intercept	0.0060 (0.0019)
Real time consumption growth	0.838 (0.092)
R^2	67.0%

Table 23: Regression of the conventional revised measure of consumption growth (Δc_t) on the corresponding real time measure (Δc_t^{RT}) over the period 1965-2008.

The real risk-free rate has been a problematic quantity to measure accurately since true real risk-free assets have not been available in the market until recent times (further, even these recent innovative instruments have limited liquidity). This quantity has therefore usually been approximated by subtracting some measure of expected inflation from the nominal risk-free rate. Two implicit assumptions are made in doing so. The first is that the measure of expected

A25

inflation being used is reasonable and the second is that inflation uncertainty risk is unpriced. This first issue is a serious one for data over long time frames since inflation survey data is typically unavailable over the entire period. The general convention in this regard has been to use lagged or realized inflation. Hence, it is important to understand the possible size of error that using lagged or realized inflation can cause. We therefore regress the measures of the real risk free rate made using lagged, realized and expected inflation on each other to estimate the amount of measurement noise. We tabulate the results in table 24. We see that the R^2 of each of the regressions is quite low with the average being under 33%. This indicates that the measurement noise in the reported real risk free rate has about twice the variance of the underlying quantity. This is in line with the size of inflation forecast errors as reported by Croushore (2010). Hence, for the simulations, we model the measured real risk free rate as the actual risk free rate plus iid noise with twice it's realized variance.

We do not investigate the second issue (the assumption of zero price of risk for inflation risk uncertainty) but note that it's inclusion will only strengthen the results in our favor as the stock portfolios are real assets and their P/D ratios are real yields and thus are not directly exposed to inflation risk.

C.4 Type I error of Asset Pricing Tests with Respect to the Long Run Risk Model

We now analyze the performance of tests of four different asset pricing restrictions of the long run risk model in order to determine which is the most reasonable one to use in the analysis in this paper.

Regression of $r_{f,t}^{\text{lagged}}$ against $r_{f,t}^{\text{realized}}$

Coefficient	Estimate (Std. Err.)
Intercept	0.0046 (0.0028)
$r_{f,t}^{\text{realized}}$	0.454 (0.106)
R^2	23.6%

Regression of $r_{f,t}^{\text{lagged}}$ against $r_{f,t}^{\text{expected}}$

Coefficient	Estimate (Std. Err.)
Intercept	-0.0023 (0.0030)
$r_{f,t}^{\text{expected}}$	0.890 (0.145)
R^2	38.6%

Regression of $r_{f,t}^{\text{realized}}$ against $r_{f,t}^{\text{expected}}$

Coefficient	Estimate (Std. Err.)
Intercept	-0.0007 (0.0035)
$r_{f,t}^{\text{expected}}$	0.859 (0.169)
R^2	30.4%

Table 24: Regression of three measures of the real risk free rate on each other. The three measures are computed using the lagged, realized and expected inflation. The regressions are restricted to the post-1946 period as expected inflation data is only available for it.

The asset pricing restrictions that we consider are related to the one analyzed by Ferson, Nallareddy, and Xie (2012).[37] It is[38]

$$E[r_{l,t+\Delta t} - r_{f,t}] + \frac{1}{2}\text{Var}[r_{l,t+\Delta t} - r_{f,t}] \approx \beta_{l,\tilde{\epsilon}}\lambda_{\tilde{\epsilon}} + \beta_{l,\tilde{w}}\lambda_{\tilde{w}} \qquad (114)$$

where the returns $r_{l,t}$ are continuously compounded, $\tilde{x}$ and $\tilde{\sigma^2}$ are the estimated values of x_t and σ_t^2 and $\tilde{\epsilon}$ and $\tilde{w}$ are the estimated values of the innovations of these processes. x and σ^2 are

---

[37]Ferson, Nallareddy, and Xie (2012) use GMM with the Euler moment restrictions in the SDF framework. We use the beta representation which is approximate but quite accurate when dealing with continuously compounded returns.

[38]Note that we don't need a $\beta_{\Delta c}$ term as there is no contemporaneous correlation between the dividend growth and consumption growth innovations in this model.

estimated in the same manner as by Bansal, Yaron, and Kiku (2007) and Ferson, Nallareddy, and Xie (2012), i.e. by the use of the following regressions

$$\Delta c_{t+\Delta t} = \alpha_0 + \alpha_1 z_{m,t} + \alpha_2 r_{f,t} + \sigma_t \eta_{t+\Delta t} \sqrt{\Delta t} \tag{115}$$

$$\tilde{x}_t = \alpha_0 - \mu + \alpha_1 z_{m,t} + \alpha_2 r_{f,t} \tag{116}$$

$$\tilde{x}_{t+\Delta t} = \rho \tilde{x}_t + \tilde{\epsilon}_{t+\Delta t} \sqrt{\Delta t} \tag{117}$$

$$\sigma_t^2 \eta_{t+\Delta t}^2 \Delta t = \beta_0 + \beta_1 z_{m,t} + \beta_2 r_{f,t} + \omega_{t+\Delta t} \tag{118}$$

$$\tilde{\sigma}_t^2 \Delta t = \beta_0 + \beta_1 z_{m,t} + \beta_2 r_{f,t} \tag{119}$$

$$\tilde{\sigma}_{t+\Delta t}^2 = \nu \tilde{\sigma}_t^2 + \tilde{w}_{t+\Delta t} \sqrt{\Delta t} \tag{120}$$

where $z_{m,t}$ is the log market P/D ratio (taken to be the log P/D ratio of the first asset in the simulations) and Δt is one year.

The alternate asset pricing restrictions that we consider are analogous but use the two largest estimated log P/D ratio factors instead of the log market P/D ratio and the real risk free rate as they should also span x and σ^2. The principal idea behind this approach is that given the null, they should be more accurately estimated in the presence of measurement error since they are estimated using multiple assets. The asset pricing restriction analogous to (114) is then given by

$$E[r_{l,t+\Delta t} - r_{f,t}] + \frac{1}{2} \mathrm{Var}[r_{l,t+\Delta t} - r_{f,t}] \approx \sum_{i=1}^{2} \beta_{l,IF_i} \lambda_{IF_i} \tag{121}$$

where F_i and IF_i are the ith principal components of the log P/D ratios of the assets and their estimated innovations respectively (the latter are estimated by fitting the former to an AR(1) process).

The above asset pricing relations do not include the $\beta_{l,\sigma^2}^2 \lambda_{\sigma^2}$ term, which, as we have shown in theorem 1, should be included in order to make it exact. Excluding it, however, has

A28

very minor impact on the results both because the short time series in the data and the high persistence of the long run risk processes makes these βs difficult to estimate precisely and also because this term captures the variance in expected excess returns which is much smaller than the expected excess returns. The latter is true because the order of magnitude of the expected excess returns is 0.1 while that of the variance of the expected excess returns is the square of this number, i.e. 0.01. Another reason we exclude this term is to ensure that we are comparing our procedure with the methodology in the extant literature.

We examine whether the hypothesis that the factors being considered are useless is rejected by the cross sectional regression methodology. This is done using the Wald test for the risk premia of the factors with their covariance matrix being estimated in the standard manner (see for eg., (Shanken 1992) and (Shanken and Zhou 2007)). The rejection frequencies for each of these tests in 1000 simulations are reported in table 25. The results show that the test of the asset pricing restriction involving the log P/D ratio factors (which also include noise calibrated to fit the observed factor structure of log P/D ratios) display much greater power than those involving the estimated long run risk processes. Hence, we use the former in our analysis in this paper.

Hypothesis	Non-rejection rate		
	p=0.10	p=0.05	p=0.01
$\lambda_{\tilde{\epsilon}}, \lambda_{\tilde{w}} = 0$	12.3%	24.0%	50.4%
$\lambda_{IF_1}, \lambda_{IF_2} = 0$	0.4%	0.6%	1.5%

Table 25: Rejection frequencies for the hypothesis that the λs of the relevant factors are zero.

C.5   Conclusion

In this appendix, we simulate a 25 asset long run risk economy with parameters chosen so as to match key economic and financial moments with those in U.S. economic and financial data. We analyze the type I error of different asset pricing tests within this economy and find, when realistic measurement noise is introduced into it, that tests using estimates of the long run

risk components derived from projections of consumption growth onto the log market price dividend ratio and real risk free rate display high type I error while those estimating the same components using the principal components of the log price dividend ratios of the assets do not do so. This implies that the latter type of tests have a more desirable profile. Hence, we use such tests in this paper.
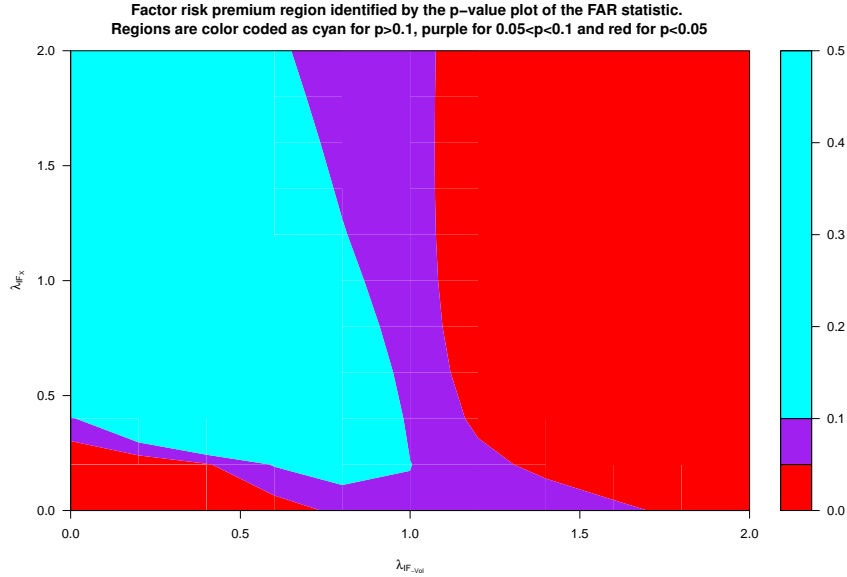
D   Robust Test Statistics

Figure 9: p-value plot of the test of the joint hypothesis of factor pricing together with $(\lambda_{IF_{-Vol}}, \lambda_{IF_X}) = (\hat{\lambda}_{IF_{-Vol}}, \hat{\lambda}_{IF_X})$ using the FAR statistic proposed by (Kleibergen 2009) for the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008. $\lambda_{IF_{-Vol}}$ and λ_{IF_X} are the factor risk premia for IF_{-Vol} and IF_X. F_{-Vol} and F_X are the negative volatility and consumption/dividend growth factors and IF_{-Vol} and IF_X are their innovations. F_{-Vol} and F_X are the negative volatility and consumption/dividend growth factors and IF_{-Vol} and IF_X are their innovations.

Since the excess returns of the 25 Fama-French portfolios formed on the basis of size and book to market ratio have a strong factor structure, it is important to use robust test statistics

to eliminate the problem of useless factors being identified as useful (a problem forcefully brought out by Kleibergen (2009) and Kleibergen (2010)). Hence, we use the robust test statistics suggested by Kleibergen (2009) to ensure that the factors here are not useless.

We find that these robust test statistics reject the joint hypothesis that $\lambda_{IF_{-Vol}} = \lambda_{IF_X} = 0$ (non-rejection of the hypothesis would indicate that the pricing factors are useless) and do not reject either the hypothesis of factor pricing or that of $\lambda_{IF_{-Vol}} = \hat{\lambda}_{IF_{-Vol}}, \lambda_{IF_X} = \hat{\lambda}_{IF_X}$ for many values of $(\hat{\lambda}_{IF_{-Vol}}, \hat{\lambda}_{IF_X})$ including those estimated using the cross sectional regressions (rejection of this would indicate that the model is rejected by the data). Figure 9 contains the plot of the p-values of the FAR test statistic for many different values of $(\hat{\lambda}_{IF_{-Vol}}, \hat{\lambda}_{IF_X})$. This statistic tests the joint hypothesis of factor pricing and of $\lambda_{IF_{-Vol}} = \hat{\lambda}_{IF_{-Vol}}, \lambda_{F_X} = \hat{\lambda}_{IF_X}$. It shows that the joint hypothesis is rejected at $\hat{\lambda}_{IF_{-Vol}} = \hat{\lambda}_{IF_X} = 0$ and also that it is not rejected for many other values of $\hat{\lambda}_{IF_{-Vol}}$ and $\hat{\lambda}_{IF_X}$ including those in table 4. Further, the region identified by $p > 0.1$ excludes $\lambda_{IF_X} = 0$ but not $\lambda_{IF_{-Vol}} = 0$. This is consistent with the findings using GMM which are analyzed in the next subsection.

The JGLS statistic which tests the hypothesis of factor pricing for a given value of $\lambda_{IF_{-Vol}}$ and λ_{IF_X} is plotted in figure 10. Since it tests a weaker hypothesis, it is not surprising that it rejects fewer values of $\lambda_{IF_{-Vol}}$ and λ_{IF_X} than the FAR statistic. When combined with the GLS-LM statistic, also plotted in figure 10, which tests the hypothesis that $\lambda_{IF_{-Vol}} = \hat{\lambda}_{IF_{-Vol}}, \lambda_{F_X} = \hat{\lambda}_{IF_X}$ given that factor pricing is correct, it gives very similar results to those given by the FAR statistic.

Hence, we can conclude that the robust test statistics show that we cannot reject the hypothesis of factor pricing with the innovations of the two identifid log P/D factors However, these resuls, together with the findings made through GMM, do cast some doubt on the significance of $\lambda_{IF_{-Vol}}$.
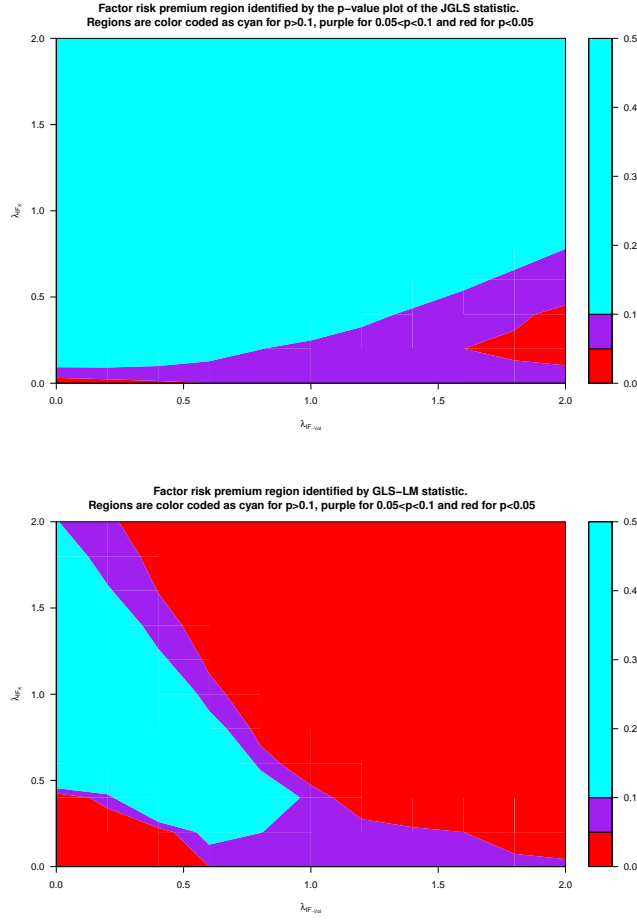
Figure 10: p-value plot of the test of the hypothesis of factor pricing given $\left(\lambda_{IF_{-Vol}}, \lambda_{IF_X}\right) = \left(\hat{\lambda}_{IF_{-Vol}}, \hat{\lambda}_{IF_X}\right)$ using the JGLS and GLS-LM statistics proposed by Kleibergen (2009) for the 25 Fama-French portfolios formed on the basis of size and book to market ratio over the period 1944-2008. $\lambda_{IF_{-Vol}}$ and λ_{IF_X} are the factor risk premia for IF_{-Vol} and IF_X. F_{-Vol} and F_X are the negative volatility and consumption/dividend growth factors and IF_{-Vol} and IF_X are their innovations. F_{-Vol} and F_X are the negative volatility and consumption/dividend growth factors and IF_{-Vol} and IF_X are their innovations.