The multi-horizon dynamics of risk and returns

JOB MARKET PAPER

Andrea Tamoni,[1] (First version: March 2010.) This version: October 5, 2011

Preliminary and Incomplete

Abstract

I study the effects of changes in risk on asset prices across different time horizons (or time-scales) and provide a new insight into the dynamics of equity premia. I find that, contrary to the implication of standard model such as the Consumption-CAPM, risk premia are weakly related to consumption volatility at short horizons whereas long-run past volatility strongly determines the long-run dynamics of expected stock returns. More importantly I show that a model specified at a fixed time-scale may not necessarily lead to obtain a significant long-term risk-returns relation upon aggregation of the one-period dynamics of volatility and returns. I thus develop a consumption-based model that simultaneously characterizes both the short- and long-term behaviors of risk and returns and successfully replicate the pattern observed in the data. Whereas previous empirical literature has mainly focused on stock market volatility, when I estimate the model I am able to relate movements of equity premia at specific frequency intervals to sources of macroeconomic risk, as measured by conditional volatility of consumption. The empirical results emphasize the importance of simultaneously modeling consumption at multiple time-scales and point to changing consumption volatility as an important long-run priced factor.

JEL Classification Codes: E32, E44, G11, G12.

Keywords: Asset Pricing, Risk-Return, Macroeconomic Uncertainty, Multi-scale time series models.

---

[1] Bocconi University; andrea.tamoni@unibocconi.it. Earlier versions of this paper were circulated under the title "Persistence Heterogeneity in Volatility and Long-Run Risk-Return Trade-Off". I would like to thank my committee, Carlo Favero, Fulvio Ortu and Claudio Tebaldi. I am also greatly indebted to David Backus, Mikhail Chernov and Stanley Zin for their helpful comments and discussions. All remaining errors are mine.

1 Introduction

I am interested in the behavior over multiple horizons of the relation between expected excess returns and risk, proxied by the conditional volatility of returns or consumption, for three main reasons: first, understanding what drives the dynamics of the equity premium, the expected excess return on a market portfolio over the risk-free interest rate, is a fundamental problem of financial economics; second, the idea that changing volatility of consumption or aggregate cash flows can affect asset prices and equity premia is intuitive and has a long-standing place in the asset pricing literature;[2] and third, time horizon (i.e. holding period) is almost as important a consideration as asset classes for investors.

With very few exceptions,[3] the great majority of the empirical literature investigating the natural hypothesis that changing risk premia are induced by movements in volatility does assume that the only relevant source of information is at the fine scale of resolution. As a point of departure from this previous body of work I argue that data observed at different time-scales reveal different information and there is the need for an asset pricing model to combine and integrate the information arising at these different levels of resolution.[4]

To get an intuition of the behavior of the risk-return trade-off at different scales, it is useful to look at Figure 2. Following the approach of Bandi and Perron (2008) I run multi-horizons regressions of future returns on past consumption volatility. The Figure shows the R^2 attained in these regressions for US, UK and Canada and provides evidence of a weak relation at short-horizon and of a strong one at long-horizon, the R^2 being increasing from around zero at the 1-year horizon to more than 50% at the 10-years horizon. These empirical results are especially interesting for two reasons. First, the statistical weak risk-return relation at short-horizon runs counter the strong intuition of a positive relation between volatility and

[2]Early work investigating the relationship between consumption volatility and expected excess returns includes Abel (1988), Barsky (1989), Giovannini (1989), Kandel and Stambaugh (1990), and Gennotte and Marsh (1992). More recently, Bansal and Yaron (2004) have taken this idea to a model of recursive preferences of the type explored by Epstein and Zin (1989, 1991), and Weil (1989), showing that a reduction in consumption volatility can raise asset prices if the intertemporal elasticity of substitution is greater than unity.

[3]See Ghysels, Santa-Clara and Valkanov (2005) and Bandi and Perron (2008).

[4]The term "time-scale" may be viewed as "resolution". At high time-scales (low frequencies, long-term) there is a coarse resolution of a time series, while at low time-scales (high frequency, short-term), there exists a fine resolution. Moving from low time-scales to high time-scales (from short-term to long-term) leads to a more coarse characterization of the time series due to averaging. In this study, high-frequency refers to variability on time-scales of one year, whereas the lowest frequency refers to decadal variations as well as trends. Table 3 shows how to interpret each scale in terms of frequency interval.

expected returns that comes from the above mentioned models. Second, and more important, I show that these results are inconsistent with general equilibrium models where the time series properties of variables driving equity returns, e.g. consumption growth, are defined only for the finest scale. In fact simple aggregation of a model for volatility and returns specified at one time horizon, for instance monthly, may not necessarily lead to obtain a long-term risk-returns. A question naturally arises. What feature are necessary to generate this counterintuitive behavior of expected returns and volatility?

In this article, I address this question by developing a new asset pricing model that successfully and simultaneously characterize both the short- and long-term behaviors of a time series. To this end, I depart from the conventional time series analysis in the way I model changing consumption volatility, moving away from specifications in which the focus is exclusively on a given time-scale and all implications at coarser levels of aggregation may be obtained by simply scaling up the series (e.g. by non-overlapping averages). Conversely I propose a new modeling approach built in a cascade way from coarse to fine scales in order to incorporate information observed at different scales of resolution. More specifically I couple standard linear models at different levels of resolution using an autoregressive process for each level to allow the existence of relevant dynamics at multiple resolution levels.

In modeling macroeconomic risk in this manner, I draw on an extensive body of work in the finance literature that studies the memory in volatility at different time intervals and finds evidence of volatility clustering at all time-scales, as well as evidence of multifractality in the moment-scaling behavior of the data. In a related paper, Calvet and Fisher (2007) capture volatility persistence across time-scales and long memory using the multifractal model of asset returns. In the same vein, Ghysels et al. (2005) study the predictability of return volatility at different frequencies by employing mixed data sampling regressions. Our paper also relates to the recent work by Sizova (2010), where the author show that it is important to account for long-range dependence in predictive variables when considering long-horizon regressions. Our multi-scale time series approach has in fact the capacity to emulate long-memory processes. However the main advantage of our approach over actual long memory models is interpretability - the long-memory type of behavior is explicitly modeled as a result of high autocorrelation in the coarse level of the hierarchy.

In order to preserve tractability and to provide insight into the relation between risk and return, our model specification is kept simple and developed in the context of a representative agent, exchange economy (Lucas, 1978). The model generates realistic features of the relation between risk and returns that are broadly consistent with the empirical evidence and delivers a number of testable implications.

The empirical part of this article follows the multiresolution approach suggested by our model and characterizes the relation between equity premia and macroeconomic risk at all time-scales. Our estimation evidence is based on the data sampled annually. We use a long span of data over the 1930-2010 period that covers a wide range of macroeconomic events that potentially contain important information about variation in consumption growth volatility. The empirical results can be summarized as follows. First, "macroeconomic uncertainty", as measured by the persistent components of consumption growth variance, is an important source of aggregate risk and an important driver of long-run expected market returns. Our work uncovers an important relation between macroeconomic uncertainty, consumption risk and asset prices and thus complements the existing literature which has mainly looked at just equity market risk, see e.g. Glosten, Jagannathan and Runkle (1993), Whitelaw (1994) and Boudoukh, Richardson and Whitelaw (1997) for the intertemporal relation between equity risk and return at short-horizon and Bandi and Perron (2008) for the long-run. Second, combining information at different levels of aggregation, we find a significantly positive relation between stock market risk and returns. In particular the estimate of risk aversion γ is around 4, which lines up well with economic intuition about a reasonable level of risk aversion. Finally, my work points to new directions for empirical work on the dynamics of interest rates and risk. In particular I found evidence for "macroeconomic uncertainty" as a key channel driving the time variation in the real short-term interest rate.

The rest of this article is organized as follows. Sections 2 and 3 document the failure of conventional asset pricing models, focusing exclusively on a time series at a given scale, in explaining the nature of the risk-return trade-offs. Section 4 introduce the principles of multiresolution analysis and the multi-scale modeling approach. Section 5 presents an asset pricing model that incorporates the multiresolution approach, provides intuition behind the risk-return relation in this setting and evaluates how well it performs in explaining the mild dependence at short horizons along with the strong statistical relationship in the long-run. I then explore the statistical relations implied by our model and document common movements in measures of the price-dividend ratio for the aggregate stock market and volatility of measured consumption growth. Section 7 concludes.

2 Macroeconomic Uncertainty and Risk-Return Trade-offs

The purpose of this section is twofold. First I illustrate how macroeconomic risk can affect asset prices by using simple models. Second I show that none of these models is quite sufficient to represent the whole structure of risk and returns reported in Figure 2.

and cannot helps us in interpreting the empirical evidence that the dependence of excess market returns on past consumption variance increases with the horizon and is strong in the long-run (i.e., between 7 and 10 years).

2.1 A standard set-up

I start by presenting a classic specification, an endowment economy as in Lucas (1978) where there is a representative agent who maximizes a time-separable power utility function given by:

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$$

with γ being the risk aversion parameter. The stochastic discount factor is equal to the intertemporal marginal rate of substitution in aggregate consumption, C_t:

$$M_{t,t+1} = \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}$$

Let the process for aggregate consumption growth $g_{t+1} \equiv \log \frac{C_{t+1}}{C_t}$ follow:

$$\begin{aligned} g_{t+1} &= \mu + v_t \eta_{t+1} \\ v_{t+1}^2 &= (1-\varphi_v)v + \varphi_v v_t^2 + \sigma_w w_{t+1} \end{aligned}$$

where v_t^2 is the conditional variance and η_{t+1}, w_{t+1} are i.i.d. $\mathrm{N}(0,1)$ shocks. Plugging the above dynamics for consumption into the expression for the log stochastic discount factor $m_{t,t+1} \equiv \log M_{t,t+1}$ yields

$$\begin{aligned} m_{t,t+1} &= \delta_1 g_{t+1} \\ &= \delta_0 + \delta_1 v_t w_{t+1} \end{aligned} \tag{1}$$

where $\delta_1 = -\gamma$. For any asset i, consider the Campbell and Shiller (1988) linear approximation for the log return

$$r_{i,t+1} \approx \kappa_0 + gd_{i,t+1} + \kappa_{i,1} pd_{i,t+1} - pd_{i,t} \tag{2}$$

where $pd_{i,t} \equiv \log\left(\frac{P_{i,t}}{D_{i,t}}\right)$ is the log price-cash flow ratio and $gd_{i,t}$ the log cash flow growth rate. To make our point and keep the discussion brief I assume that dividends are perfectly correlated with consumption and thus I will focus on the asset valuation associated with the claim to the consumption stream.

The price-dividend ratio (once it has been solved endogenously) is an affine function of the state variable, i.e.

$$pd_{i,t} = b_0 + b_\sigma v_t^2$$

Given the solution for the price-dividend ratio it is possible to derive the innovation to the asset return can as a function of the evolution of the state variables and the parameters of the model:

$$r_{t+1} - E_t[r_{t+1}] = v_t \eta_{t+1} + b_\sigma \kappa_1 \sigma_w w_{t+1} \tag{3}$$

As asset returns and the pricing kernel in this model economy are conditionally log-normal, the risk premium for any asset is determined by the conditional covariance between the return and $m_{t,t+1}$:

$$E_t[r_{i,t+1} - r_{f,t}] = -\text{cov}_t(m_{t,t+1}, r_{i,t+1}) - 0.5\text{Var}_t(r_{i,t+1}) \tag{4}$$

Exploiting the innovations in (1) and (3) into (4) it follows that:

$$E_t[r_{t+1} - r_{f,t}] = \gamma v_t^2 - 0.5\text{Var}_t(r_{t+1}) \tag{5}$$

i.e. expected returns are determined by the conditional volatility of aggregate consumption.

Of course, this stylized model has important limitations, but its very simplicity serves to illustrate the crucial point: macroeconomic risk plays an important role in determining asset values. Moreover, as shown in the next section, the relation (5) is not specific to this set-up but in fact can be obtained in a variety of models.

2.2 Economic Uncertainty and Recursive Preferences

I now present an economic model adapted from Bansal and Yaron (2004) and I show that adopting recursive preferences together with a channel for time varying expected growth does not alter the main conclusion of the previous section, i.e. expected returns are determined by the conditional volatility of aggregate consumption.

Consider an endowment economy as in Lucas (1978) where the representative agent has Epstein and Zin (1989) - Weil (1990) preferences. In this economy the intertemporal marginal rate of substitution is

$$M_{t+1} = \exp\left(\theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1}\right)$$

where the parameter ψ, is the intertemporal elasticity of substitution (IES), and $\theta = (1 - \gamma)/(1 - (1/\psi))$. The return, $r_{c,t+1}$ denotes the log return on wealth, which in this economy is simply the return on the claim to the consumption stream. Let the process for aggregate consumption growth,

$$\begin{aligned} g_{t+1} &= \mu + x_t + v_t \eta_{t+1} \\ x_{t+1} &= \rho x_t + \phi_e v_t e_{t+1} \\ v_{t+1}^2 &= (1 - \varphi_v) v + \varphi_v v_t^2 + \sigma_w w_{t+1} \end{aligned}$$

where x_t is the conditional expected growth rate and $e_{t+1}, \eta_{t+1}, w_{t+1}$ are i.i.d. $\mathrm{N}(0, 1)$ shocks.

Again to make our point and keep the discussion brief I will focus on the asset valuation associated with the claim to the consumption stream. Exploiting the Euler equation for valuing any asset $r_{i,t+1}$

$$E_t\left[\exp\left(\theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1} + r_{i,t+1}\right)\right]$$

the solution for the log priceconsumption ratio is $pc_t = b_0 + b_x x_t + b_\sigma$ where

$$\begin{aligned} b_x &= \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \varphi_v} \\ b_\sigma &= \frac{0.5[(\theta - \frac{\theta}{\psi})^2 + (\theta b_x \kappa_1 \phi_e)^2]}{\theta(1 - \kappa_1 \varphi_v)} \end{aligned}$$

In the model with Epstein-Zin preferences the risk premium takes the following expression:[5]

$$E_t[r_{t+1} - r_{f,t}] = \gamma v_t^2 + \lambda_e \kappa_1 b_x \phi_e v_t^2 + \lambda_w \kappa_1 b_\sigma \sigma_w^2 - 0.5 \mathrm{Var}_t(r_{t+1}) \quad (6)$$

[5]Details of the derivation are available from the author upon request. See also Bansal and Yaron (2004) and Bansal, Yaron and Kiku (2009).

where $\lambda_e \equiv (1-\theta)\kappa_1 b_x \phi_e$ is the market price of expected growth rate risk, $\lambda_w \equiv (1-\theta)\kappa_1 b_\sigma$ is the market price of volatility risk. The first term in the premium is the familiar i.i.d. case where risk aversion multiplies consumption volatility. The second term captures the exposure of the asset return to expected growth rate news. The third term is the compensation for risk associated with fluctuating consumption volatility and is absent in the case of power utility (when $\theta = 1$).

The above discussion implies that, although in the model with Epstein-Zin preferences volatility shocks carry a separate risk premium, it still holds that risk premia is a linear affine function of consumption volatility, that is

$$E_t[r_{t+1} - r_{f,t}] = \gamma_0 + \gamma_1 v_t^2 \tag{7}$$

which has the same form as the equation for the risk premia asserted in (5).[6]

3 Short- and Long-Run Implications of Fixed Time-Scale Model

In the previous section I showed that in a variety of models,[7] expected returns are determined by the conditional volatility of aggregate consumption, that is:

$$E_t[r_{t+1} - r_{f,t}] = \gamma_0 + \gamma_1 v_t^2 \tag{8}$$

$$v_{t+1}^2 = (1-\varphi_v)v + \varphi_v v_t^2 + \sigma_w w_{t+1} \tag{9}$$

where the second equation models the variance process as a persistent autoregressive process.

As noted in Hansen (2008), when I build dynamic economic models, I typically specify transitional dynamics over a unit of time for discrete-time models or an instant of time for continuous time models. Long-run implications are encoded in such specifications. I therefore consider the implication of equations (8) and (9) upon an aggregation scheme where both the equity premium and the variance are measured over the same horizon h. More formally

[6] Note that if the IES is greater than one then it immediately follows that b_x is positive. In addition, if γ is also greater than one (i.e., $\theta < 0$, λ_e, and hence γ_1, is positive.

[7] Additional models not considered in the previous section are the ones in Hansen and Singleton (1982, 1983), Abel (1988), Kandel and Stambaugh (1990) and Campbell (1993).

consider:

$$\underbrace{E_t[r^{ex}_{t+1} + r^{ex}_{t+2} + \ldots + r^{ex}_{t+h}]}_{\text{expected return at scale } h} = \beta \underbrace{(v^2_t + v^2_{t-1} + \ldots + v^2_{t-h-1})}_{\text{Risk at scale } h} \tag{10}$$
$$v^2_{t+1} = (1-\varphi_v)v + \varphi_v v^2_t + \sigma_w w_{t+1}$$

where $r^{ex}_{t+1} \equiv r_{t+1} - r_{f,t}$ and we think of the right-hand side of (10) as a measure of the uncertainty over horizon h.[8] Figure 1 shows on simulated data from a long-run risk economy that the R^2 from a regression of the equity premium on the risk, as measured by the consumption variance would be high at short horizon and small at long-horizons.

[Insert Figure 1 about here.]

I next bring relation (8) and its aggregated counterpart (10) to the data. In order to do so I need to find a proxy both for expected returns and consumption volatility. Following many empirical studies I employ realized excess returns to proxy for the equity premium. With regard to economic uncertainty we consider three empirical measures for consumption volatility; the first two specifications are not parametric whereas the last one is. The first measure is obtained computing standard deviations over h-year (past) rolling windows. The second measure, motivated by Andersen, Bollerslev and Diebold (2003a) and used in Bansal, Khatchatrian and Yaron (2005) is the realized volatility. We begin by fitting an AR(1) process for consumption growth:

$$g_{t+1} = \mu + \beta g_t + u_{t+1}$$

Then we calculate h-period realized volatility as the sum of the absolute values of the residuals over K periods:

$$v_{t-h,t} \equiv \log \sum_{i=0}^{h-1} (|\eta_t|)$$

[8] Indeed if we assume that consumption growth is unpredictable then we have that

$$\begin{aligned}\mathrm{Var}_t(g_{t,t+h}) &= \mathrm{Var}_t\left(\mu h + \sum_{i=0}^{h-1} v_{t+i}\eta_{t+i+1}\right)\\ &= \sum_{i=0}^{h-1} v^2_{t+i}\end{aligned}$$

Taking logs makes the volatility measure less sensitive to outliers and does not qualitatively affect any of our empirical results. Our third specification for volatility is parametric and is based on modeling consumption growth as following an AR(1)-HARCH(5). HARCH(n) is a variant of the ARCH model when we consider the arrival of heterogeneous information and it was introduced by Muller, Dacorogna, Dave, Olsen, Pictet and von Weizsacker (1997). As discussed below, our results are robust to these alternative measures of volatility. Appendix A contains a detailed description of the data. In all, to evaluate the relation between return and risk at different horizons h I consider regressions of the type:

$$r_{t,t+h} = \alpha_h + \beta_h \text{Risk}_{t-h,t} + \epsilon_{t,t+h} \qquad h = 1, \dots, 10 \tag{11}$$

where $r_{t,t+h}$ denotes excess market returns between years t and $t+h$, $\text{Risk}_{t-h,t}$ denotes past consumption variance, and $\epsilon_{t,t+h}$ is a forecast error. The relationship between returns and variance across different levels of aggregation h is what I call the multi-scale behavior of the risk-return trade-off.

Table 1 report the results of this set of regressions. The results in the first column show that when returns and variance are measured at the highest frequency possible[9] there is only weak evidence that periods of high stock market/consumption volatility coincide with periods of predictably high stock returns. The empirical evidence that expected stock returns are weakly related to volatility appears to contradict the model implication that risk and return are positively related. This is consistent with previous estimates of the relation between risk and return which often have been insignificant and sometimes even negative, see Lettau and Ludvigson (2010) for a thorough discussion.[10] Table 1 shows that although the dependence between excess market returns and past consumption variance is statistically mild at short horizons (thereby leading to a hard-to-detect risk-return trade-off, as in the existing literature) it increases with the horizon and is strong in the long-run (i.e., between 7 and 10 years), consistent with the recent empirical findings in Bandi and Perron (2008). For aggregation levels $h = 7, 8, 9$, and 10 years the former relation reveals a pronounced dependence as highlighted by a monotonously increasing R^2 that peaks to 55% at the 10-year horizon. Importantly this result is insensitive to how volatility is measured.

[9]Tests of the short-run risk-return trade-off have been mainly conducted with returns computed for a holding period of one month. For our long data sample 1929-2010 consumption data are available at annual frequencies and fix the highest frequencies time interval for our empirical analysis.

[10]French, Schwert and Stambaugh (1987), Baillie and DeGennaro (1990) and Campbell and Hentschel (1992) do find a positive albeit mostly insignificant relation between the conditional variance and the conditional expected return. In contrast, Campbell (1987) and Nelson (1991) find a significantly negative relation. Glosten et al. (1993), Harvey (2001), and Turner et al. (1989) find both a positive and a negative relation depending on the method used.

[Insert Table 1 about here.]

Two natural concerns however arise with regard to the above results. Our work analyzes one US based dataset, with one history of returns and variance. Using this data set alone, it is hard to definitively exclude the possibility of spurious results. International data offer an interesting test of our hypothesis of a multi-scale behavior of the risk-return trade-off. Thus I estimate the horizon-dependent relation(s) (11) between the conditional mean of excess market returns and variance of consumption growth for different countries and I plot in Figure 2 the estimated R^2. As follows from the figure, the R^2 derived in regressions of future returns on past volatility is increasing from around zero at the 3-month horizon to more than 60% for US, UK and Canada. This results support a positive long-run risk-return relation in international markets.

[Insert Figure 2 about here.]

The second concern[11] regards the use of the sum of squared daily returns, i.e. the second moment, as a proxy for market risk. This holds true when the drift of returns is negligible and in particular for very short horizons. The long-horizon results could simply reflect the effect of the drift on the second moment in order to deliver a constant variance. To ensure the robustness of our results, I thus consider a measure of realized volatility used by Bansal et al. (2005) that account explicitly the time-varying mean of consumption growth at long-horizons. Results are reported in Panel B of Table 1 and are broadly consistent with the one obtained using volatility from the AR(1)-HARCH(5) specification.

Overall whereas much work has been devoted to assessing the validity of the classical short-run ($h = 1$) risk-return trade-off I have shown that the models testable restrictions are rejected on a stronger ground when the risk-return relation is studied across different horizons. This is consistent with the results in Bandi and Perron (2008) and Sizova (2010). These works explore the long-run implications of traditional short-term risk-return models such as (8) augmented with a classical (autoregressive) process for variance as (9), and prove that simple temporal aggregation of such short-term specification cannot imply the long-term results.

[11] I thank Francesco Corielli for pointing this out.

The above facts reveal that the dynamics of risk and returns change with the time horizon and the risk-return trade-off should be considered time-scale dependent. The interval size of the time grid on which volatility is measured is therefore an essential parameter. Volatility measured with high resolution contains information that is not covered by low resolution volatility and vice versa.

The failure of the previous models can be attributed to the fact that they typically model only one time-scale (usually a month) at time. In the next section I make sure that, indeed, this is the case and even expanding the state space by allowing the volatility process to be the sum of two conventional models, where one has nearly a unit root, and the other has a much more rapid decay does not solve the empirical puzzle as far as these components are specified at the same time-scale.

3.1 Too big to be true

As follows from the pattern shown in Figure 2 the R^2 for the return regression seems very high. In this section I conjecture that the observed patterns in Figure 2 emerged spuriously. My first goal is to check if the type of model described by equations (8) and (9) can capture the long-run predictability pattern. Assuming a nearly integrated framework as in Valkanov (2003), Bandi and Perron (2008, Proposition 3) show that, under the assumption that $\lim_{T\to\infty} h/T = \lambda$ (where T is the sample size) the R^2 in regression(s) (11) converges to a nondegenerate random variable:[12]

$$R^2 \Rightarrow \frac{\left(\int_{\lambda}^{1-\lambda} A(\tau)B(\tau)d\tau\right)^2}{\left(\int_{\lambda}^{1-\lambda} A^2(\tau)d\tau\right)\left(\int_{\lambda}^{1-\lambda} B^2(\tau)d\tau\right)} \tag{12}$$

$$\tag{13}$$

where $A(\tau)$ and $B(\tau)$ are processes on the interval $[0,1]$ whose expressions are given in equations (B.5) and (B.6) and the symbol $\Rightarrow$ denotes weak convergence as $T \to \infty$. I can now study the implications of a short-memory model that takes into account overlapping observations, persistence in the predictive variable and the effect of a negative correlation between shocks to returns and shocks to variances. To examine the asymptotic distribution we consider reasonable values for the coefficients of the processes $A(\tau)$ and $B(\tau)$. Appendix

[12]See Appendix B for a proof.

B describes our choice. Using formula (12), I construct a distribution for R^2 in regression(s) (11) when the DGP is given by the Fixed Time-Scale Model, see equations (8) and (9).

[Insert Table 2 about here.]

The results are reported in Table 2 for the horizon $H = 10$ and $T = 81$ observations, as in our sample. Table 2 shows that the the limiting distribution of the R^2 is more concentrated around zero. This is contrary to our findings. However the resulting values of R^2 are very imprecisely measured. For example, in Case 1, the 90% probability interval for R^2 stretches from 0.0% to 25.7%. Note that despite the high dispersion of estimates, a value of R^2 from regression(s) (11) above 50% is rarely observed. As follows from the data in Table 2, the probability of this event is equal to 0.1-0.4% for all cases.

Overall these observation leads to the conclusion that fixed-time scale models have hard time in in revealing and assessing long-run predictability in returns based on volatility.

3.2 Expanding the State-Space

Intuition could lead to think that a model where volatility is made up of the sum of independent autoregressions, each with a different persistence level can be better suited to capture short-term and long-term effects such as the ones presented in the previous section. Superposition models of this type have become popular in financial econometrics as they are more general than empirically limiting Markov volatility models while close to corresponding continuous time models (Engle and Lee (1993); Barndorff-Nielsen and Shephard, 2001; Chernov, Gallant, Ghysels and Tauchen (2003)). In this section I instead show that expanding the state space cannot help us in interpreting the empirical evidence presented in the previous section. In particular I assume a two-factor stochastic volatility model for both the consumption and dividend claim:

$$g_{t,t+1} = \mu + \sqrt{dv_{p,t}^2 + (1-d)v_{s,t}^2}\eta_{t+1} \tag{14}$$

$$gd_{t,t+1} = \mu_d + \varphi_d\sqrt{cv_{p,t}^2 + (1-c)v_{s,t}^2}u_{t+1} \tag{15}$$

where $v_{p,t}^2$ and $v_{s,t}^2$ are the slowly evolving component and the quickly mean reverting one, respectively and u_{t+1} is correlated with w_{t+1}. The dynamics of the two fundamental uncertainty factors are:

$$\begin{aligned} v_{p,t+1}^2 &= (1-\varphi_{v,p})v + \varphi_{v,p}v_{p,t}^2 + \sigma_p w_{p,t+1} \\ v_{s,t+1}^2 &= (1-\varphi_{v,s})v + \varphi_{v,s}v_{s,t}^2 + \sigma_s w_{s,t+1} \end{aligned} \tag{16}$$

If I consider an endowment economy with Epstein and Zin (1989) - Weil (1990) preferences, then the log stochastic discount factor innovations are:

$$\begin{aligned} m_{t,t+1} - E_t[m_{t,t+1}] &= \left(-\frac{\theta}{\psi} + \theta - 1\right)\sqrt{dv_{p,t}^2 + (1-d)v_{s,t}^2}\eta_{t+1} + \dots \\ &\quad + (\theta-1)\kappa_1 b_{\sigma,p}\sigma_{p,w}w_{p,t+1} + (\theta-1)\kappa_1 b_{\sigma,s}\sigma_{s,w}w_{s,t+1} \end{aligned} \tag{17}$$

and the stock market innovation is

$$\begin{aligned} r_{t+1} - E_t[r_{t+1}] &= \mu_d + \varphi_d\sqrt{cv_{p,t}^2 + (1-c)v_{s,t}^2}u_{t+1} + \dots \\ &\quad + \underbrace{\kappa_{1,m}b_{\sigma,p}^m}_{\beta_{p,mw}}\sigma_{p,w}w_{p,t+1} + \underbrace{\kappa_{1,m}b_{\sigma,s}^m}_{\beta_{s,mw}}\sigma_{s,w}w_{s,t+1} \end{aligned} \tag{18}$$

Equation 18 shows that the implied stock return process must also have the same two volatility factors.

Plugging (17) and (18) into (4) I finally obtain:

$$E_t[r_{t+1} - r_{f,t}] = \gamma(dv_{p,t}^2 + (1-d)v_{s,t}^2) + \lambda_{p,w}\beta_{p,mw}\sigma_{p,w}^2 + \lambda_{s,w}\beta_{s,mw}\sigma_{s,w}^2 - 0.5\mathrm{Var}_t(r_{t+1})$$

where $\lambda_{i,w} \equiv (1-\theta)\kappa_1 b_{\sigma,i}$ for $i = \{p, s\}$. Note that once again I obtain a relation similar to (7) where risk premia is a linear affine function of consumption volatility. The interesting question now is to see what role do the two volatility components play in our new model? Can the superposition model given by 16 replicate the risk-return empirical pattern reported in Figure 2 ?

To answer these questions I simulate from an economy calibrated following the model of Zhou and Zhu (2010), which extend the Bansal and Yaron (2004) and Bansal et al. (2009) models by introducing both a long-run and a short-run volatility component into the consumption and dividend processes. Figure 1 shows that the superposition model cannot reconcile the findings of a weak short-run and a sizable long-run risk-return trade-off, and has indeed the same problem facing the one-factor volatility model presented in Section 2.

Conventional time series analysis, focusing exclusively on a time series at a given scale, lacks the ability to explain the nature of the data-generating process. The problem of aggregation arises: the properties of long-horizon risk-return cannot be covered by a model equation that focuses on a series of volatility and returns at the highest frequency of observation. The new challenge to theoreticians is the development of consistent models that successfully and simultaneously characterize both the short and long-term behaviors of a time series.

Since fixed time-scale are not adequate for capturing the perception of risk and return, I argue that a better insight into the dynamics of financial markets can be achieved with a time-adaptive framework that simultaneously takes all time-scales of the statistical process into account, the Multiresolution analysis.

In the next section I thus quickly introduce the Multiresolution Approximation framework to analyze a time series on different scales, with different degree of resolution, simultaneously. I then re-investigate the dynamics of returns using a new asset pricing model that incorporates the multiresolution technique and show that the incorporation of multiple time-scales into the analysis should improve the interpretation of the empirical evidence just presented.

4 The concept of a time series with more than one time scale.

Multiresolution decomposition. A time series could contain one characteristic time scale, or it could contain processes at several different resolutions or time-scales. Using multiresolution analysis I can separate the different time-scales in a given time series.

There are a number of reasons why we might want to decompose a time series into its component scales. First, the signal being investigated may possess features and significant dynamics at several different scales. Second, the process may live at one scale and the noise another. If the signal can be separated from the noise at a particular time scale, the hope is to arrive at a more predictable representation of the signal (see for instance Ortu, Tamoni and Tebaldi (2011b) for an application to long-run risk in consumption).

4.1 An insight into the time series decomposition

The concept of a multiresolution analysis (MRA) is that a given time series, with finite variance, may be decomposed into different approximations associated with unique resolutions (or time horizons). Next I briefly explain this idea since I shall use it for the new asset pricing model throughout the paper. For a complete and formal treatment of the multiresolution analysis, I refer to Mallat (1989), Dijkerman and Mazumdar (1994) and Levan and Kubrusly (2006).

A multiresolution representation provides a simple hierarchical framework for interpreting the time series information, see Ortu, Tamoni and Tebaldi (2011a). Given a sequence of increasing resolutions $\{r_j\}_{j \in Z}$, I wish to build a multiresolution representation based on the differences of information available at two successive resolutions r_J and r_{J+1}. The details of a time series at the resolution r_j are defined as the difference of information between its approximation at the resolution r_j and its approximation at the lower resolution r_{j-1}. Following the approach of Burt et al. (1983) and Crowley (1987) I choose to work with dyadic (or pyramidal) scales so that our resolution levels will be given by $r_j = 2^j$ where $j \in Z$.

More formally let A_{2^j}, be the operator which approximates a stochastic process with paths in L^2 at a resolution 2^j. Here, I characterize A_{2^j}, from the intuitive properties that one would expect from such an approximation operator. I state each property in words, and then give the equivalent mathematical formulation.

1. A_{2^j} is a linear operator. If $A_{2^j}x$ is the approximation of some signal x at the resolution 2^j, then $A_{2^j}x$ is not modified if I approximate it again at the resolution 2^j. This principle shows that $A_{2^j} \circ A_{2^j} = A_{2^j}$. The operator A, is thus a projection operator on a particular vector space $V_{2^j} \in L^2$, i.e. $A_{2^j} = \text{proj}_{V_{2^j}}$. The vector space V_{2^j}, can be interpreted as the set of all possible approximations at the resolution 2^j of signals in L^2.

2. Among all the approximated functions at the resolution 2^j, $A_{2^j}x$ is the function which is the most similar to x.

$$\forall y \in V_{2^j}, \|y - x\| \geq \|A_{2^j}x - x\|$$

   Hence the operator A_{2^j} is an orthogonal projection on the vector space V_{2^j}.

3. The approximation of a signal at a resolution 2^j contains all the necessary information to compute the same signal at a coarse resolution 2^{j+1}. This is a causality property. Since A_{2^j}, is a projection operator on V_{2^j}, this principle is equivalent to

$$\forall j \in \mathbb{Z} \; V_{2^j} \subset V_{2^{j-1}}$$

I call any set of spaces $\{V_{2^j}\}_{j \in m\mathbb{Z}}$ which satisfies the properties (2)-(3) (additionally also certain self-similarity relations in time and scale must be satisfied, see Mallat (1989)) a multiresolution approximation of L^2. The associated set of operators A_{2^j}, satisfying (l)-(3) give the approximation of any L^2 signal at a resolution 2^j. We saw that the approximation operator A_{2^j} is an orthogonal projection on the vector space V_{2^j}. In practice, any stochastic signal can be observed only at a finite resolution. For normalization purposes, I suppose that this resolution is equal to 1.

In order to numerically characterize this operator, I must find an orthonormal basis of V_{2^j}. Meyer (1988) shows the existence of such a basis (or multiresolution filter) and that these bases generalize the Haar basis. Let's call this orthonormal basis h.

Let $A_1 x$ be the discrete approximation at the resolution 1 that is measured. The causality principle says that from $A_{2^j} x$ I can compute all the discrete approximations $A_{2^j} x$ for $j \geq 1$. I now describe a simple iterative algorithm for calculating these discrete approximations. Let's call $\pi_t^{(j)}$ the approximated signal obtained from $A_{2^j} x$. The next coefficients $\pi_t^{(j)}$ are generated from the scale coefficients $\pi_t^{(j-1)}$ by convolving the latter with the low-pass filter h:

$$\pi_t^{(j+1)} = \sum_{l=-\infty}^{+\infty} h(l) \pi_{t+2^j l}^{(j)} \tag{19}$$

Since $V_{2^j} \subset V_{2^{j-1}}$, $\pi_t^{(j+1)}$ is a coarser approximation of x than is $\pi_t^{(j+1)}$. Therefore, the key idea of MRA consists in studying a signal by examining its coarser and coarser approximations by canceling more and more details from the data.

MRA allows now to build a representation of the series based on the differences of information available at two successive resolutions 2^j and 2^{j+1}. I now explain how to extract the difference of information between successive resolutions. The information that is removed when going from one approximation to the next coarser one is called the detail. The detail coefficients at scale $j + 1$ (i.e. those coefficients whose information content is the difference between consecutive approximations, say at levels j and $j+1$ in the decomposition) are defined formally as:[13]

$$x_t^{(j+1)} = \pi_t^{(j)} - \pi_t^{(j+1)}$$

Hence the detail coefficients can also be viewed as differences between weighted averages where the weights are determined by a given multiresolution filter.

A multiresolution analysis thus allows to rewrite the information in the original series g_t as a collection of details at different resolutions and a low-resolution approximation:[14]

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_t^{(J)} \tag{20}$$

Therefore, the time series should be viewed as a cumulative sum of the detail variations $x_t^{(j)}$ and will become smoother as j increases. The detail coefficients $\{x_t^{(j)}\}_{j=0}^{J}$ decompose the information from the original time series into pieces associated with both time and scales. In particular the level j detail coefficients $\{x_t^{(j)}\}_t$ are associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and the scaling coefficients $\{\pi_t^{(J)}\}_t$ are associated with averages on a scale of length 2^J. Since the detail coefficients capture the variation of the time series at a given scale and interval of time, our approach is to model the detail coefficients directly. Importantly, at any time point, t, I never use information (time-wise) after k in calculating the wavelet coefficient, thus the algorithm produces an adapted (non anticipative) decomposition.

I propose multiresolution stochastic models on the discrete wavelet coefficients as approximations to the original time process.

4.2 Haar multiresolution and temporal aggregation

Infinitely many MRAs exist (see for example Aldroubi and Unser, 1993; Abry and Aldroubi, 1995). In the following I select the Haar only to underline the possibility of exactly reformulating the aggregation procedure as an MRA.[15] Aggregating the data means averaging

[13] Note that $x_t^{(j+1)}(\omega)$ is a random variable. A sample $x_t^{(j+1)}(\omega)$ is the detail coefficient at (j,t) of a sample trajectory $x_t(\omega)$.

[14] Reconstruction of the process from its discrete detail coefficients is understood in the L^2 sense. See also Ortu et al. (2011a).

[15] More generally any MRA can therefore be understood as an aggregation procedure. See Abry, Veitch and Flandrin (1998).

them over a time duration T; in other words, it means filtering the data a low-pass function whose characteristic time support is of length T.

The Haar multiresolution is designed from the following simple filter $h = \left(\frac{1}{2}, \frac{1}{2}\right)$. In this case equation (19) becomes

$$\pi_t^{(j+1)} = \frac{1}{2}[\pi_t^{(j)} + \pi_{t-2^j}^{(j)}]$$

I now study the decomposition (20) for level $j = 1$. First consider convolving the input data x_t with the filter h. This yields

$$\begin{aligned}
\pi_t^{(1)} &= \frac{1}{2}[\pi_t^{(0)} + \pi_{t-1}^{(0)}] \\
&= \frac{x_t + x_{t-1}}{2} \\
x_t^{(1)} &= \pi_t^{(0)} - \pi_t^{(1)} \\
&= \frac{x_t - x_{t-1}}{2}
\end{aligned}$$

where I define $\pi_t^{(0)} = x_t$.

From this definition, it is straightforward to check that the aggregated process $\{x_{t,t+H}\}_t$ where $x_{t-H,t} = \sum_{k=0}^{H-1} x_{t-k}/H$ with aggregation level $H = 2^h$, can be rewritten as an approximation of x:

$$\begin{aligned}
x_{t-2^{(h-1)},t} &= \pi_t^{(h)} \\
&= \sum_{j=h}^{J} x_t^{(j)} + \pi_t^{(J)}
\end{aligned} \tag{21}$$

This relation answer the question: what is happening at different scales in the series (scale behavior of a series). Moreover it shows that there is an exact, obvious and natural identity between the aggregation procedure and Haar MRA: studying x over longer and longer observation periods T simply translates in the MRA vocabulary to increasing the scale of analysis 2^j, or equivalently to lowering the resolution.

4.3 Inspecting the Mechanism

Before introducing the full-blown asset pricing model built on the multiresolution analysis presented, I now sketch a multi-scale time series model that can be used to interpret the results in Figure 2.

Based on the multiresolution framework presented in the previous section and using the decomposition (20) and the aggregation relation (21) I build the dynamics of the volatility process from coarse to fine level of resolution:

$$\begin{aligned} v_{t,t+h} &= v_{t+h}^{(p)} && \text{at coarse scale } h \gg 1 \\ v_{t,t+1} &= v_{t+1}^{(p)} + v_{t+1}^{(s)} && \text{at the finest scale} \end{aligned}$$

and similarly for the returns

$$\begin{aligned} r_{t,t+h} &= r_{t+h}^{(p)} && \text{at coarse scale } h \gg 1 \\ r_{t,t+1} &= r_{t+1}^{(p)} + r_{t+1}^{(s)} && \text{at the finest scale} \end{aligned}$$

Consider again the regression(s) in (11):

$$\begin{aligned} \underbrace{r_{t,t+1}}_{r_{t+1}^{(p)}+r_{t+1}^{(s)}} &= \alpha_1 + \beta_1 \underbrace{v_{t-1,t}^2}_{v_t^{(p)}+v_t^{(s)}} + \epsilon_{t+1} \\ \underbrace{r_{t,t+h}}_{r_{t+h}^{(p)}} &= \alpha_h + \beta_h \underbrace{v_{t-h,t}^2}_{v_t^{(p)}} + \epsilon_{t+h} \end{aligned}$$

It is now easy to see that if I interpret the decomposition of both the volatility and the return as the sum of an information component, captured by the long-run processes $r_{t+1}^{(p)}$ and $v_{t+1}^{(p)}$, and a noisy one, captured by $r_{t+1}^{(s)}$ and $v_{t+1}^{(s)}$, then an error-in-variable problem will lower, under certain conditions,[16] the R^2 at short horizon compared to the (true) one at long horizon. Thus the proposed method that decomposes a given time series on a scale-by-scale basis, based on multiresolution approach, can allow to interpret the fact that relationship between the return and risk becomes stronger as the scale increases. In the next section I extend the simple set-up above to a model which consider the multi-scale nature of risk and return at all possible horizons.

5 A new multi-horizon asset pricing model

Following the approach discussed in Section 4, I incorporate in the standard long-run risk model the decomposition of time series into components realized over different time horizons,

[16] See Lancaster (1963).

see (20), so that the log consumption growth, g_t takes the following form:

$$g_t = \sum_{j=1}^{J} g_t^{(j)}$$

where $\{g_t^{(j)}\}_j$ are meant to capture the behavior of the original time series over time-scale of length 2^{j-1} periods. In fact the aggregation relation (21) assures that, for sufficiently high $h = 2^{J-1}$ the dynamics of consumption growth at horizon h are fully determined by its persistent component:

$$\begin{aligned} g_{t,t+h} &\equiv \log \frac{C_{t+h}}{C_t} \\ &= \sum_{k=0}^{h-1} g_{t+h-k}/h \\ &= g_{t+h}^{(J)} \end{aligned}$$

This says that a process equation that successfully explains the horizon h, 10 year consumption changes, for example, is unable to characterize the nature of yearly consumption changes $g_{t,t+1}$.

Now, I closely follow the approach of Bollerslev, Tauchen and Zhou (2009) and Zhou (2010) and suppose that the growth rate of consumption in the economy is not predictable. The novelty is to assume that each component of consumption growth, $g_{j,t}$ is driven by its own stochastic volatility, $v_{j,t}$, i.e.:

$$\begin{aligned} g_{t+2^j}^{(j)} &= v_{j,t}\ e_{j,t+2^j}^{g} \\ e_{j,t+2^j}^{g} &\sim N(0,1) \end{aligned} \tag{22}$$

with the shocks $e_{j,t+2^j}^{g}$ being uncorrelated. To close the dynamics of the model I assume that each of the stochastic variance components $\left\{v_{j,t}^2\right\}_{j=1}^{J}$ is observed over time intervals of different sizes and follows its own autoregressive process, i.e.

$$\begin{aligned} v_{j,t+2^j}^2 &= \rho_j v_{j,t}^2 + \varepsilon_{j,t+2^j} \\ \varepsilon_{j,t+2^j} &\sim N\left(0, \left(\sigma^{(j)}\right)^2\right) \end{aligned} \tag{23}$$

This statistical description of volatility is called multi-scale autoregressive process, for it has an autoregressive process occurring on all scales $j = 1, \ldots, J$. Moreover using equations (22) together with the decomposition (D.2) I can compute the volatilities realized over different time horizons:

$$\begin{aligned} \mathrm{Var}_t[g_{t+1}] &= v_{p,t}^2 + v_{s,t}^2 \\ \mathrm{Var}_t[g_{t,t+h}] &= v_{p,t}^2 \end{aligned} \tag{24}$$

This show that *at each time-scale* the variance is described by a different autoregressive structure. This is the key difference between the superposition model suggested by Barndorff-Nielsen and Shephard (2001) and the multiresolution approach. In the superposition model, volatility *at the lowest scale* is the sum of independent autoregressive processes, each with a different persistence level. The multiresolution approach takes instead advantage of the persistence *across scales*. My model has heterogeneity derived from the fact that different autoregressive structures are present at each time scale allowing volatility components to decay at different rates. With this regard I sometimes refer to j as the level of persistence of the j-th components. Importantly, equations (22) to (23) represent a natural way to incorporate persistence heterogeneity in the macroeconomic uncertainty framework while retaining its pedagogical simplicity.

Finally note that in this simple specification of the model variance can go negative. To ensure the positivity of the variance process, one can pursue the approach of Barndorff-Nielsen and Shephard (2001) and assume that the innovations in the components of variance processes follow a Gamma distribution. This would just slightly complicate the algebra without adding much to the intuition to the model. In the following I therefore assume that consumption variance shocks are Gaussian to explain the major implications of the model. However when I estimate the consumption variance components I impose positivity by estimating the log-variance.

To give economic and structural meaning to the parameters I assume, as in BY04, a pure exchange economy with a representative agent with Epstein-Zin recursive preferences. The well known Euler condition for such an agent is:

$$E_t\left[e^{m_{t+1}+r^i_{t+1}}\right] = 1 \tag{25}$$

where m_{t+1} is the log stochastic discount factor given by

$$m_{t+1} = \theta \log\beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r^a_{t+1} \ , \tag{26}$$

r^a_{t+1} is the log return of the claim which distributes a dividend equals to aggregate consumption.[17] The parameter β is the preference discount factor. The preference parameter ψ measures the intertemporal elasticity of substitution, γ measures the risk aversion and $\theta = (1-\gamma)\,/\,(1-1/\psi)$.

In what follows I provide the basic steps to determine the pricing kernel and risk premia on the market portfolio in my long-run risk model with persistence heterogeneity.[18] Recall first that by the standard Campbell and Shiller (1988) log-linear approximation for returns one obtains:

$$r_{a,t+1} = \kappa_0 + \kappa_1 z^a_{t+1} - z^a_t + g_{t+1} \tag{27}$$

where z^a_t denotes the log price-consumption and the log price-dividend ratio respectively. Recalling the decomposition of consumption into components with different levels of persistence, and denoting with $z^{(j)}_{a,t}$ the components with persistence j of the (log) price-consumption ratio, it is natural to conjecture that there exists *component by component* a linear relation between the financial ratios and the state variables $v^2_{j,t}$, i.e.

$$z^{(j)}_{a,t} = A_{0,j} + A_j v^2_{j,t} \tag{28}$$

As long as A_j these relations tells us that the variation in valuation ratios can be attributed to fluctuations in economic uncertainty. These relations together with equation 23 imply that measures of economic uncertainty (conditional variance of consumption) are predicted by components of valuation ratios.

The values of $A_{0,j}, A_j$ in terms of the parameters of the model are obtained from the Euler condition (25) after the log stochastic discount factors and the returns are all expressed in terms of the factors $\{v^2_{j,t}\}_{j=1}^J$ and of the innovations $\{e^g_{j,t+2^j}\}_j$ and $\{\varepsilon_{j,t+2^j}\}_j$. In Appendix D I show that plugging these expressions for the stochastic discount factor and for the returns into the Euler equation and using the method of undetermined coefficients one obtains a set of equations for the coefficients $A_{0,j}, A_j$, the solution of which are given by the following vectors of sensitivities:

$$\underline{A} \;=\; 0.5\frac{\left(\theta - \frac{\theta}{\psi}\right)^2}{\theta}\left(\mathbb{I}_J - \kappa_1 M\right)^{-1}\underline{1}$$

[17] To make my point and keep the discussion brief I will focus on the asset valuation associated with the claim to the consumption stream. In the previous version of this paper we show how to price a claim to dividends which is modeled as a levered claim on the consumption components processes containing additional independent shocks. Nonetheless, the general structure for the asset risk premium and its valuation ratio is analogous to the one presented in this section.

[18] All details behind our calculations are given in the Appendix D.

where

$$M = diag\,(\rho_1, \ldots, \rho_J)$$

and $\underline{A}$ denotes the column vectors with entries, $A_1, \ldots, A_J$. Two features of this model specification are noteworthy. First, if the IES and risk aversion are larger than 1, then θ is negative, and a rise in volatility lowers the price-consumption ratio. Similarly, an increase in economic uncertainty will make consumption more volatile, which lowers asset valuations and increases the risk premia on all assets. This highlights that an IES larger than 1 is critical for capturing the negative correlation between price-dividend ratios and consumption variance. Second, an increase in the permanence of variance shocks, that is M, magnifies the effects of volatility shocks on valuation ratios, as changes in economic uncertainty are perceived as being long-lasting.

To study the scale-dependent consequences of my model for the equity premium in Appendix I show that innovation in the returns' component at level of persistence j are given by:

$$r^{(j)}_{a,t+2^j} - E_t[r^{(j)}_{a,t+2^j}] = v_{j,t}e^g_{j,t+2^j} + \kappa_1 \left(A_j \varepsilon_{j,t+2^j}\right)$$

and that the innovations of the stochastic discount factor's components are given by

$$m^{(j)}_{t+2^j} - E_t[m^{(j)}_{t+2^j}] = -\lambda_g v_{j,t}e^g_{j,t+2^j} - \lambda_j \varepsilon_{j,t+2^j} \quad j = 1, \ldots, J \tag{29}$$

Recall that the risk premium on any asset i satisfy, in this set-up, $E_t[r_{i,t,t+h} - r_{f,t+h}] + 0.5\sigma^2_{r_{i,t,t+h}} = -\text{cov}_t(m_{t,t+h}, r_{i,t,t+h})$ where $r_{i,t,t+h}$ and $m_{t,t+h}$ are the stock return and stochastic discount factor aggregated over h-period. With the innovations to the equilibrium returns at hand and using (29) together with (21) one finally can compute the risk premia for the consumption claim asset, $r_{a,t+2^{h-1}}$ and for the market portfolio, $r_{m,t+2^{h-1}}$ for any horizon 2^{h-1}:

$$\begin{aligned} E_t[r_{a,t,t+2^{h-1}} - r_{f,t+2^{h-1}}] + 0.5\sigma^2_{r_{a,t,t+2^{h-1}}} &= -\text{cov}_t\left(\sum_{j=h}^{J} m^{(j)}_{t+2^j}, \sum_{j=h}^{J} r^{(j)}_{a,t+2^j}\right) \\ &= \lambda_g \sum_{j=h}^{J} v^2_{j,t} + \kappa_1 \underline{\lambda}_\varepsilon \mathbf{Q} \underline{A}' \end{aligned} \tag{30}$$

where $\lambda_g \equiv \left(\frac{\theta}{\psi} - \theta + 1\right)$, $\quad \underline{\lambda}_\varepsilon \equiv \kappa_1(1-\theta)\underline{A}$ and $\mathbf{Q} = \mathbf{E}_t\left[\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}'_{t+1}\right]$.

This model, however simple, provides a natural platform on which to investigate the risk behavior at different time horizons. This is done by accounting for changes in the covariance at different scales. Intuitively, equity markets price not only high-frequency but also diverse lower-frequency fundamentals such as demographics (Abel, 2003), technological innovation (Pastor and Veronesi, 2005), and variations in cash-flows and macroeconomic uncertainty (Bansal and Yaron, 2004; Lettau, Ludvigson, and Wachter, 2004) as in my specific case.

6 Empirical evidence

In this section I provide new evidence that relates asset prices to consumption and stock market variance. Appendix A contains a detailed description of the data.

Along this section I extract consumption variance using an AR(1)-HARCH(5). Importantly this estimation technique accounts for changes in the volatility dynamics at different scales. The filtered time-varying consumption growth volatility is persistent with an autocorrelation coefficient of about 0.75. Nevertheless as I have argued in the previous sections, the volatility itself can be interpreted as the sum of different components each one realized over different time horizons indexed by j. I extract such components using the Haar MRA, see equations (19) and (20). I instead measure stock market risk by the realized variance obtained using high-frequency (i.e. daily in this case) return data and then I decompose it using the Haar MRA, see equations (19) and (20). For an interpretation of the cycle durations corresponding to the time-scale level j in the case of annual times series see Table 2.

6.1 Long-run Risk and Return Trade-off

In this Section I discuss the risk-return trade-off implied by my model. Whereas the body of empirical evidence on the risk-return relation is mixed and inconclusive, as discussed in the introduction, here I argue that the disagreement in the empirical literature on the risk-return relation is likely to be attributable to not properly considering the different volatility components associated with different horizons.

Equation (30) yields a set of relation at a fixed level of persistence j between the component of expected market returns and the component of macroeconomic uncertainty, $v_{j,t}^2$:[19]

$$E_t[r_{m,t+2^j}^{(j)} - r_{f,t+2^j}^{(j)}] = \lambda_g v_{j,t}^2 + const \tag{31}$$

[19] If one sums over j the relation (31) and applies the forward decomposition (D.2) to the right-hand side then (30) is retrieved.

Applying the persistence based decomposition described in Section 4 to the realized excess returns[20] and to the consumption variance, I can finally estimate the following set of regressions:

$$r^{(j)}_{t+2^j} - r^{(j)}_{f,t+2^j} = \alpha_j + \beta_j v^2_{j,t} + \epsilon_{j,t+2^j} \quad j = 1, \ldots, J$$

The results are reported in Table 4. Although all regressions are run with an intercept, I do not report their point estimate since I always find insignificance of the intercept coefficients.[21] Note that the relation is not statistically significant at medium horizons, i.e. $j = 2, 3$. However the results suggest that the (macroeconomic) risk-return relation holds true for the components at very high time-scales, in particular for $j = 4, 5$ corresponding to horizons of $8 - 16$ and $16 - 32$ years, respectively.

To further illustrate the behavior of the common cyclical components between stock market returns and consumption variance I plot in Figure 3 the medium-frequency variation in returns and consumption variance. The medium-term frequency are obtained as the sum of the components of the respective series at time-scales $j = 4, 5$. The comovement is visibly striking. The coefficient of determination from a regression of the (medium-term cycle) in returns on the (medium-term cycle in) consumption variance is $R^2 = 0.71$ and the corrected t-statistics of the coefficient loading on the risk measure is 4.43. In Section 6.4 I show that these consumption variance components indeed proxy for real economic uncertainty such as long-run unemployment and/or productivity risk. Since in my model stock market volatility is proportional to consumption volatility, see equation (D.8), I should expect market returns to comove at the medium frequency not only with consumption risk but also with market risk. This is indeed the case as shown in Figure 3 where I also plot the medium-frequency variation in stock market variance. The coefficient of determination from a regression of the (medium-term cycle) in returns on the (medium-term cycle in) market variance is $R^2 = 0.41$ and the corrected t-statistics of the coefficient loading on the risk measure is 2.83. It is important to stress that the common feature of these oscillations is that they occur over a longer time frame than is typically considered in conventional business cycle analysis. The results in Schwert (1989), Schwert (1989) and Corradi, Mele and Distaso (2008) show how difficult it is to explain low frequency fluctuations in stock market volatility through low frequency variations in the volatility of other macroeconomic variables. This is not

[20] Although the theoretical risk-return relation is based on the expected (components of) excess return, following many empirical studies I employ (components of) realized excess returns to proxy for the latent variables.

[21] I did not consider regressions for which the intercept α_j is constrained to be 0 although from a statistical standpoint, provided the restriction is true, the slope estimator is still estimated consistently but with increased precision.

surprising since these works focus on business cycle frequency and therefore tend to sweep these oscillations into the trend, thereby removing them from the analysis. My new empirical evidence, and theoretical model, support the view that stock market volatility is related to the volatility of macroeconomic variables, such as consumption and productivity as shown in Section 6.4, although not precisely related to the business cycle.

[Insert Table 4 about here.]

[Insert Figure 3 about here.]

Some remarks are in order. First of all the above evidence supports a strong relation of future long-run excess market returns with past long-run market variance. This is not in contrast with the existing literature which documents a hard-to-detect risk-return trade-off. In fact consistent with this findings the risk-return dependence is statistically mild at short horizons. Moreover these results are in line with those in Bandi and Perron (2008) who find a strongly significant correlation between $r_{t,t+h}$ and $v^2_{t-h,t}$ for values of h between 6 and 10 years. Importantly I provide a long-run model which support these results. In fact as noted in Bandi and Perron (2008) simple aggregation of short-term risk-return models under a classical (autoregressive) process for variance cannot imply these results. Therefore whereas traditional short-term risk-return models yield counter factual long-run implications my long-run risk model with persistence heterogeneity is free of this unsupported empirical implication. This is due to the fact that I do not impose a single autoregressive process for the aggregate variance observed at the highest frequency of observation but instead I use a multiscale autoregressive process to describe the variance components at each different level of persistence.

Second my methodology used to study the (long-run) risk-return trade-off can be compared to the MIDAS framework of Ghysels et al. (2005), Ghysels, Sinko and Valkanov (2007) and to the Heterogeneous Autoregressive (HAR) model proposed by Corsi (2009). In fact all these methodology share the common idea to construct regressions combining data with different sampling frequencies. However my approach differs in many respects. First MIDAS[22] exploits

[22]MIDAS come in different form, e.g. MIDAS regressions with polynomials (Ghysels, Santa-Clara and Valkanov, 2003a) and MIDAS with stepfunctions (Forsberg and Ghysels, 2004) whose HAR Model is a special case. However the basic ideas stay the same.

high frequency (financial) data to predict low frequency (macro) data whereas my approach uses time series sampled at the same frequency, for instance quarterly. Second whereas the MIDAS approach estimates the variance using a weighted average of past daily squared returns I instead allow the variance to have components with different decay rates. I then run regressions of the components of the regressand with a specific decay rate onto the components of regressors with the very same decay. This allows us to investigate whether or not one volatility component is more important than total volatility in driving the dynamics of the equity premium. My work is therefore closer to the component-GARCH approach of Engle and Lee (1993) and its extension in Maheu and McCurdy (2007).

Finally the risk-return relation is involving conditional expected return and conditional expected risk. Previous research takes diverse approaches in measuring the expected return (e.g. Campello, Chen and Zhang (2008) and Pastor, Sinha and Swaminathan (2008)) and conditional variance (e.g., French et al. (1987) and Ghysels et al. (2005)). Importantly Ludvigson and Ng (2007) note that the estimated risk-return relation is likely to be highly dependent on the particular conditioning variables analyzed in any given empirical study.[23] Here I adopt a simple approach and use average realized excess returns as a proxy for expected equity returns[24] and then extract their components using the forward decomposition. I do so because I want to stress the importance of using persistence-consistent measures of returns and variance in investigating the intertemporal risk-return relation. My results has highlighted that in order to detect a positive risk-return it is important to make sure that the measures of both the expected return and conditional variance of returns have the same level of persistence.[25]

6.2 The risk aversion

In Appendix D I show that the following relation at a fixed level of persistence j between the component of expected market returns on the aggregate consumption claim and conditional

[23]Ludvigson and Ng (2007) discuss one potential remedy to this problem based on the methodology of dynamic factor analysis for large data sets, whereby a large amount of economic information can be summarized by a few estimated factors.

[24]This practice is justified on grounds that for sufficiently long horizons, the average return will "catch up and match" expected return on equity securities and relies on a belief that information surprises tend to cancel out over the period of the study. Thus ex post average excess equity returns provide for an easy-to-implement, seemingly unbiased estimate of expected equity risk premium.

[25]Alternatively in Merton's theoretical specification it is implicitly embedded that measures of both the expected excess return and conditional variance of returns are based on the same information set. Basically my procedure define "common information set" as "the information set where all the variables have the same level of persistence"

consumption variance holds true:

$$E_t[r^{(j)}_{a,t+2^j} - r_{f,t+2^j}] = \lambda_g v^2_{j,t} + \kappa_1[\underline{\lambda}_\varepsilon \mathbf{Q}]_j A_j \quad j = 1, \dots, J$$

This is a constrained set of relations where the parameter γ, that links the information content of the excess return components to the consumption growth variance ones, is common across persistence levels.

Since in my model consumption variance drives market risk, see equation (D.8), I can obtain a more familiar risk-return relation. Indeed in my model

$$\sigma^2_{a,j,t} = v^2_{j,t} \tag{32}$$

Substituting (32) into (31) one finally obtains the risk-returns trade-off disaggregated across levels of persistence:[26]

$$E_t[r^{(j)}_{m,t+2^j} - r^{(j)}_{f,t+2^j}] = \gamma\sigma^2_{a,j,t} + const \tag{33}$$

In this special case where the parameter to be estimated is equal at all levels of persistence, one can solve the overlapping problem by adopting the technique suggested in Fadili and Bullmore (2002). In particular Fadili and Bullmore (2002) suggest to (sub)sample the components at level of persistence j with frequency 2^j[27] in order to get rid of the autocorrelation problem and then to apply to the so obtained sampled time series the generalized least squares estimator (GLS). [28]

To study the risk-return relation I need to proxy for the sample path variation in observed market returns. Based on Corollary 1 of Andersen, Bollerslev, Diebold and Labys (2003b), I assume that the conditional expectation of annual quadratic variation (QV_t) is equal to the conditional variance of annual returns, that is $\sigma^2_{a,t} = E_t[QV_{t+1}]$.[29] Assuming that the realized

[26]If one sums over j the relation (33) and applies the forward decomposition (D.2) to the right-hand side then one obtains back (30).

[27]Note that if I apply the decomposition to a time series with $T = 2^J$ elements I then obtain J components with T elements. If I subsample the components I obtain a new time series with $T/2 + T/4 + ... + T/2^J = T$ elements, that is the new sampled series has the same length of the original one.

[28]More precisely this estimator makes use of the decimated (not-redundant) Haar transform which yields a diagonalized covariance matrix of the regression errors, i.e. the off-diagonal elements can be set to zero. Diagonalization simplifies numerical identification of parameter estimates and implies that the WLS estimator is theoretically approximate to the best linear unbiased (BLU) estimator and can provide maximum likelihood estimates of both signal and noise parameters, namely γ and σ_η.

[29]Given a process $X(t)$ and the partition $T = \{t_0, \dots, t_n\}$ of $[0, t]$, the quantity $V^2(T) = \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2$ is computed and, if $\|T\| = \max_{1\le k\le n} |t_k - t_{k-1}| \to 0$ the limit is the quadratic variation (QV) of $X(t)$, i.e. $\lim_{\|T\|\to 0} V^2(T) = \langle X \rangle_t$.

variance is a unbiased estimator of quadratic variation it follows that $\sigma^2_{a,t} = E_t[RV_{t+1}]$. In the following I both use the realized variance RV_t as a proxy for $\sigma^2_{a,t}$ following the approach of Bandi and Perron (2008) and Sizova (2010) and the forecast $E_t[RV_{t+1}]$ from an ARMA(1,1) process following approaches used in early studies of variation.[30]

I report the results obtained using this approach in Table 5

[Insert Table 5 about here.]

The estimated ICAPM coefficient γ is 2.08 in the full sample, with a significant t-statistic. Most important, the magnitude of γ lines up well with the theory. According to the ICAPM, γ is the coefficient of relative risk aversion of the representative investor and a risk aversion coefficient of 4.08 matches a variety of empirical studies (see Hall (1988), and references therein). The significance of γ is robust in the subsamples, with t-statistics always higher than 2. Importantly when the returns and (consumption) risk are disaggregated across levels of persistence, the intercept α is always insignificant. The estimated magnitude and significance of the risk aversion coefficient are remarkable in light of the ambiguity of previous results.

6.3 Consumption Risk and Asset Prices

In this section I highlight additional empirical predictions of the model. In particular I analyze the underlying sources of risks that are driving asset prices by testing the relations in (28) which tell us that the component of economic uncertainty at level of persistence j should explain the corresponding component of the asset valuation ratios. I thus run the following projections:[31]

$$pd_{j,t} = \beta_0 + \beta_j v^2_{j,t} + \epsilon_{j,t} \quad j = 1, \dots, J$$

where β_j should provide an estimate of A^m_j according to relations (28). Table 6 provides the estimates, t-statistics, and R^2 from the above componentwise regressions where the columns stand for different levels of persistence $j = 1, \dots, J = 7$. I observe that the estimates for

[30] Drechsler and Yaron (2011) find that a parsimonious projection on the lagged VIX and index realized variance achieves better performance compared to a simple ARMA(1,1). I do not pursue this approach due to the short sample for which we have VIX data.

[31] Using the components of consumption volatility instead of variance does not qualitatively affect any of the empirical results.

$j = 5, 6, 7$ have significant robust t-statistics at the 5% level and the R^2 rises to 50% and 59% at level of persistence of $j = 6$ and $j = 7$ respectively. Hence the three components of consumption variance responsible for explaining the fluctuations in the price-dividend are very persistent with an half-life of $4, 8, 16$ years, respectively. These results endorse the argument of Lettau, Ludvigson and Wachter (2008) where the authors argue that the increase in asset valuation ratios is not well described as a sudden jump upward, but instead occurs gradually over several years due to (close to) permanent fluctuations in macroeconomic volatility of the order of up to 30 years. Whereas the 7-th component of consumption variance has a similar half-life,[32] my estimation technique uncovers other fluctuations in volatility, namely the ones at level $j = 5$ and $j = 6$ which will turn out to be important both for the risk-return trade-off and for the real risk-free rate variation. Importantly the signs of these relations are, as predicted by my economic model, negative. This evidence thus shows that a rise in economic uncertainty leads to a fall in asset prices. This is important because it highlights an often discussed but not verified view that aggregate economic uncertainty (i.e., real aggregate consumption volatility) has sizable effects on asset valuations and that financial markets dislike economic uncertainty.

[Insert Table 6 about here.]

In all the results in Tables 6 lead to the conclusion that the long-run components at levels of persistence $j = 6, 7$ of the current price-dividend ratio embody useful information reflecting the macroeconomic uncertainty which in turn is useful for predicting the future stock market volatility. Overall this evidence suggests that fundamental measure of macroeconomic uncertainty, as captured by the persistent components of consumption variance, play an important role in determining asset prices, especially if the perceived macroeconomic uncertainty unravels slowly.

6.4 Identification of long-run macroeconomic uncertainty

In this section I argue that those components of consumption variance that in the previous section I found to influence asset prices are indeed proxying for macroeconomic risks.

[32] In Lettau et al. (2008) the estimated low volatility state reached in the 1990s is expected to last about 125 quarters, over 30 years. The estimated 7-th component has an half-life of 16 years and therefore the full cycle can last up to 32 years.

A natural question is now what the consumption variance really stands for. Empirically, this is a latent variables to be estimated, and it is difficult to link them to known macroeconomic risks. I therefore try to give single name to the long-run economic uncertainty by linking it to macroeconomic variables. In particular I follow the lead of Malkhozov and Shamloo (2010) and Benigno, Ricci and Surico (2010). Malkhozov and Shamloo (2010) show that shocks to variance of productivity create movements in consumption variance and Benigno et al. (2010) in turn show that, in a model of the labor market with asymmetric real-wage rigidities, movements in the variance of productivity growth influence the trend of unemployment. I therefore investigate whether consumption variance is reflecting movements in long-run unemployment and/or the variance of productivity.

I obtain the time-series for the U.S. long-run mean of unemployment and the variance of productivity by computing averages and variances over h-year rolling windows, where the window length has been chosen in such a way to match the persistence of the consumption variance series. I then plot in Figures 4 the series of consumption variance along with the variance of productivity growth for the long sample 1930-2010 of U.S. data. The comovement between the two series is striking. Analogous patterns of comovement is found between consumption variance and the trend in unemployment. The correlation between long-run unemployment and consumption variance is $\rho = 0.84$ and the correlation between long-run productivity variance and consumption variance $\rho = 0.86$ (which increases to 0.88 in the the postwar period 1946-2010). Important low frequencies movements emerges in these series. In particular it is apparent the Great Moderation in the variance of productivity growth which coincides with a sharp fall in the unemployment trend and a decrease in consumption variance. This evidence confirm the very interesting feature of the data that there is a strong positive association between long-run unemployment, the variance of productivity growth and the consumption variance.

[Insert Figure 4 about here.]

In this section I have argued that the components of consumption variance that are reflected in the financial ratios do indeed proxy for macroeconomic uncertainty. Moreover I suggest the variance of productivity growth and the long-run mean of U.S. unemployment as significant determinants of this uncertainty. Combining these results with those in Section 66.3, we can conclude that financial market dislikes economic uncertainty and, given that such uncertainty comoves positively with long-run productivity uncertainty, a shift to higher productivity variance coincides with a fall in the level of share prices.

6.5 Risk-Free Rate and Macroeconomic Uncertainty

In this Section I show that my multi-horizon asset pricing model has important implications for the relation between macroeconomic uncertainty and the risk-free rate. In Appendix D.3 I derive the following expression for the risk-free rate

$$r_{f,t+1} = \beta_0 + \lambda_\eta^2 + \frac{(1-\theta)}{\theta}(\lambda_\eta)\sum_{j=1}^{J} v_{j,t}^2 \tag{34}$$

The above relation entails that real interest rates reflect consumption variance. This is in line with the simple general equilibrium model of Barsky (1989) where increased risk, proxied by consumption variability, lowers the riskless interest rate.

In order to bring the above relation (34) to the data, I first apply the decomposition (20) to the right-hand side and then I run a set of regressions of the components of the real short-term rate onto the components of consumption variance. Importantly, as suggested by the model, I restrict the coefficients loading on the variance components to be the same across levels of persistence. The results are reported in Table 7 and suggest that the risk-free rate compensates for fluctuations in risk. In the following I am going to argue that these results are mainly driven by the comovement between the very same components of consumption variance that are reflected in asset prices, i.e. those at levels of persistence the $j = 4, 5$ and the corresponding ones in the risk-free rate.

[Insert Table 7 about here.]

To start with I plot in Figure 5 the (aggregate) real risk-free rate and (minus) the approximation of consumption variance at level of persistence $j = 4$, i.e. the $h = 2^{4-1} = 8$ years fluctuations in consumption variance. The correlation between these two series is 0.20. This figure already highlights important comovements in postwar U.S. data between the (aggregate) interest rates and those components of consumption variance that, in Section 6.4, I argued to proxy for macroeconomic risk.

[Insert Figures 5, 6 about here.]

However I recall that the (constrained) component-by-component relation between the risk-free rate and the consumption variance, see (34), suggests to look not at the aggregate series but instead at the risk-free rate components that correspond to the same levels of persistence as the ones that proxy for the macroeconomic risk, i.e. $j = 4, 5$. With this regard my approach complements the one of Atkeson and Kehoe (2008) where the authors decompose the observed postwar U.S. history of nominal interest rates into a secular and a business cycle component. These components of the short-rate are intended to capture, respectively, the random walk movements in the Fed's inflation target and (response to) exogenous changes in real risk. Their model in particular suggest that the business cycle component of the interest rates should move one for one with risk. Instead my approach says that the components with persistence of 4 years of the interest rate should reflect, not the aggregate, but instead the corresponding component of macroeconomic risk. I therefore undertake the following exercise. I plot in Figure 6 the business cycle component of the risk-free rate suggested by Atkeson and Kehoe (2008) together with approximation of (minus) consumption variance at level of persistence $j = 4$. I use their approximation of the risk-free rate business cycle component (extracted using principal components analysis) because by doing so I am able to test both the fact that the comovements between the two series should become more apparent once they are filtered at the same time scale and also the robustness of my technique to alternative filtering choice. Indeed I observe that the correlation now rises to 0.30 which bring further support to the thesis that over horizon of 8 years, great part of the movements in the short-rate come from movements in conditional variances.

The above results are particularly important for two reasons. First they potentially explain the findings of Canzoneri, Cumby and Diba (2007) who document that in those models imposing that the conditional variances of the variables that enter the Euler equation are constant, the Euler equation itself does a poor job of capturing the link between the short-rate and the economy at business cycle frequencies. Since I just document that all of the movements in the short-rate come from movements in conditional variances and not from conditional means, the failure of the Euler equation is not surprising. Second they suggest to view the central bank's policy changes, namely the short-rate, as primarily intended to compensate for exogenous business cycle fluctuations in risk. Clearly as noted also in Atkeson and Kehoe (2008), this view differs substantially from the standard view, often summarized by a Taylor rule, where risk plays no role and, instead, the Fed's policy is a function of its forecasts of economic variables such as expected real growth and expected inflation.

7 Conclusion

In this paper, I show that standard economic models specified at one time-scale have a hard time in explaining the behavior of the risk-return trade-offs across different time horizons. Therefore there is the need to allow models for representing returns dynamics on multiple time horizons simultaneously.

To explain the observed pattern of the risk-return trade-off across investment horizons, i.e. consumption risk appears in the long-run to be a major contributing factor with a much smaller role when adopting a short-run perspective, I design and solve a consumption-based asset pricing model where*at each time-scale* the components of consumption variance are described by a different autoregressive structure. Our disaggregated economic models delivers the interesting empirical finding of a positive and significant long-run dependence between expected excess market returns and past variance without necessarily implying a positive short-horizon risk-return. Moreover I find a positive and significant estimate of the risk aversion coefficient once I explicitly account for the behavior of both risk and returns across different frequencies of observation.

I also obtain new results about the link between asset prices and macroeconomic uncertainty. I show that long horizons components of economic uncertainty, as measured by the persistent components of consumption variance, sharply explain valuation ratios. In particular asset valuations drop as economic uncertainty rises that is, financial markets dislike economic uncertainty. Finally I show that long-run macroeconomic volatility exerts a significant effect on the short-term interest rate. In all I conclude that once the channels associated with fluctuating economic uncertainty and economic growth are disaggregated across levels of persistence, they become important for a reasonable interpretation of asset markets.

The paper offers several possible directions for further research. First, our results have important implications regarding the scaling behavior of volatility, and also for the calculation of risk at different time horizons. Intuitively one would expect the incorporation of multiple time-scales into the analysis to improve the efficiency of risk management which requires scaling a risk measure (standard deviation, say) of one time-scale to another. Second it would be interesting and extremely informative to apply the theoretical environment of the model to the cross section of returns to explain the findings of Parker and Julliard (2005) and Bandi, Garcia, Lioui and Perron (2010) that the contemporaneous consumption or market risk explains little of the variation in average returns across the 25 Fama-French portfolios, but that a measure of consumption or market risk at a horizon of three to five years explains a large fraction of this variation.

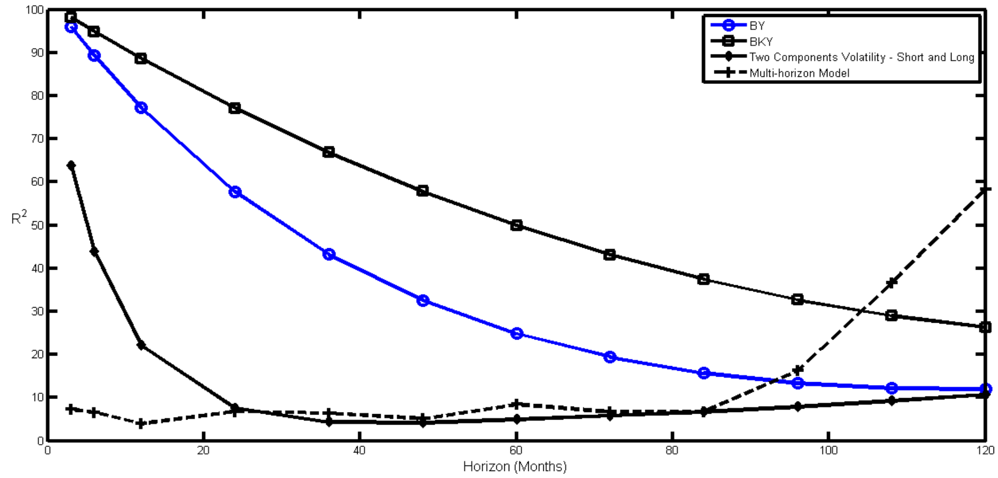

Figure 1: The Figure shows the R^2 attained at different horizons from a regression of the equity premium on the consumption variance. Dashed and solid lines denote the results obtained for the fixed time-scale model as calibrated by Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2007a). The two volatility components model is calibrated following Zhou and Zhu (2010).

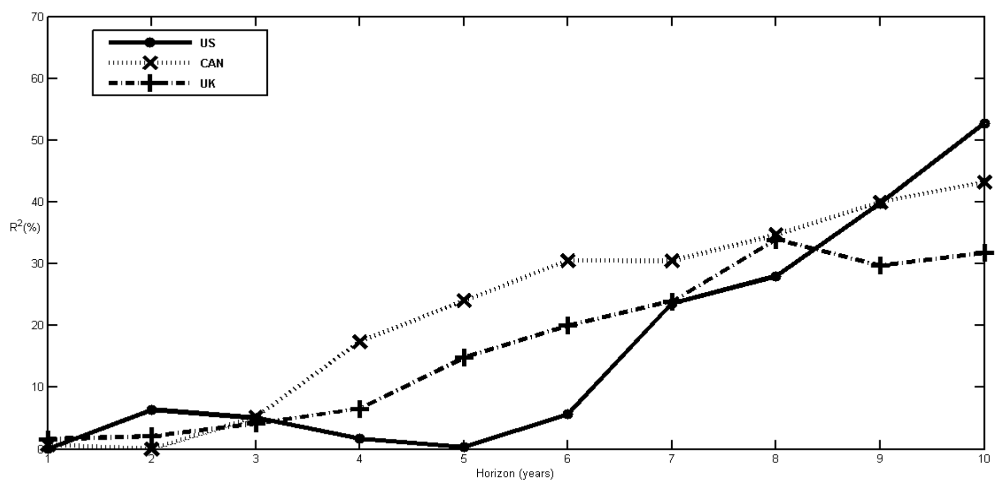

Figure 2: Multi-country Risk-Return Trade-Off. We run linear regressions (with an intercept) of h-period continuously compounded excess market returns $r_{t,t+h} = \sum_{i=1}^{h} r_{t+i-1,t+i}$ on h-period past consumption variances $v^2_{t-h,t}$. The Figure shows the R^2 for such a regression for different countries. The sample period spans from 1930 to 2009 for Canada and the UK, and from 1930 to 2010 for the US.

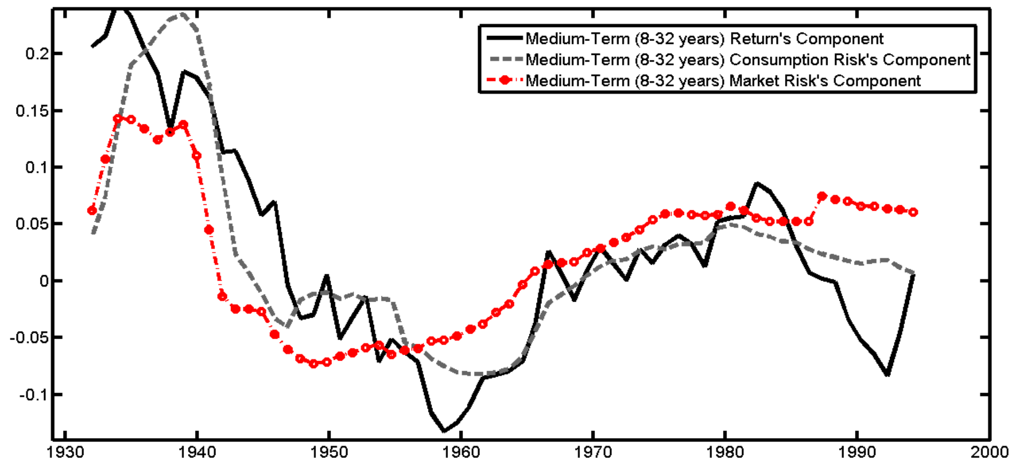

Figure 3: The Figure plots the medium-frequency component (corresponding to time-scales $j = 5, 6$, i.e. cycles of length 8-32 years) of the future compounded log returns (solid line) and the fitted value from a regression of the medium-frequency returns components on the medium-frequency past consumption variance (dashed line) and realized stock market variance (dot-dashed line). The timing is as follows: if you invested one dollar on a given date, it tells you how much total return you would have made over the following eight years. Returns are annualized. The sample spans the period 1930-2010.

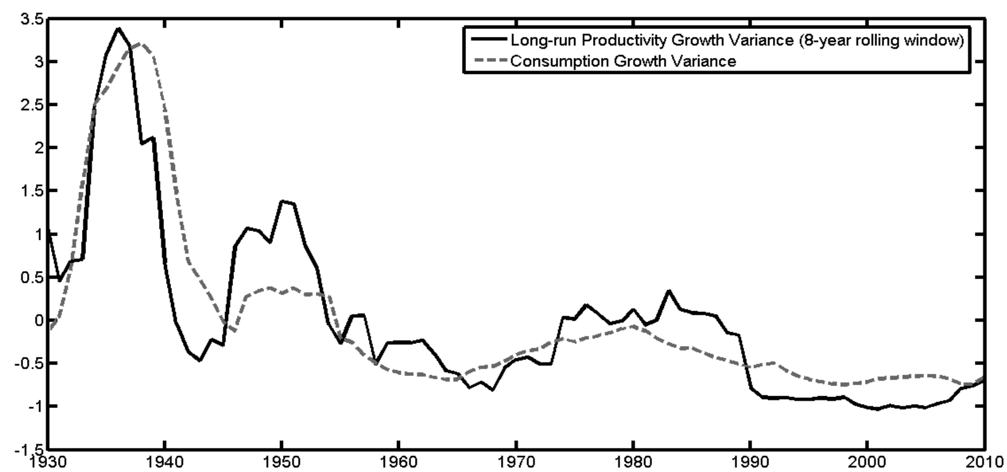

Figure 4: Approximation at level of persistence $j = 4$ of log consumption growth variance σ_t and long-run variance of productivity growth computed using eight-year rolling windows. Both series are demeaned and divided by their standard deviation.

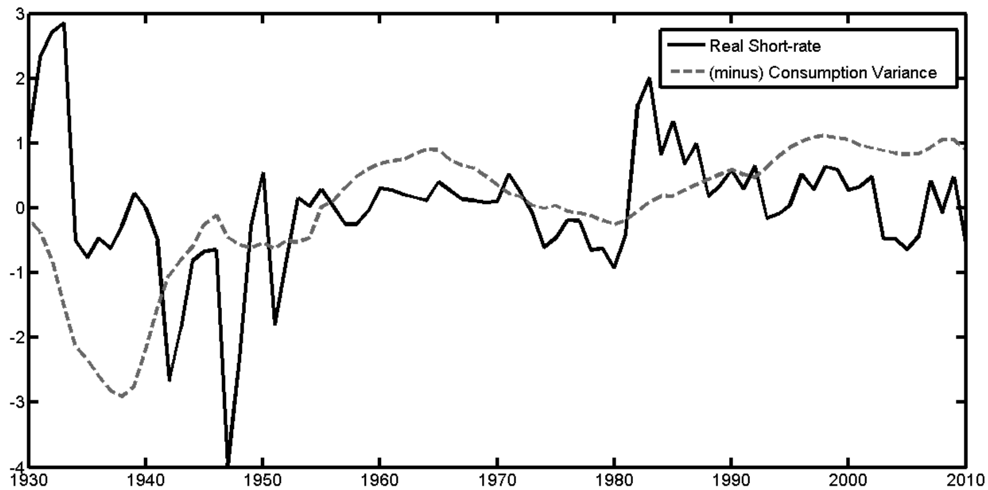

Figure 5: Short-rate and consumption variance approximated to level of persistence $j = 4$. The real return to short-term nominal investments is the 1-year treasury yield minus actual realized inflation. Both series are demeaned and divided by their standard deviation. The sample spans the period 1930-2010.

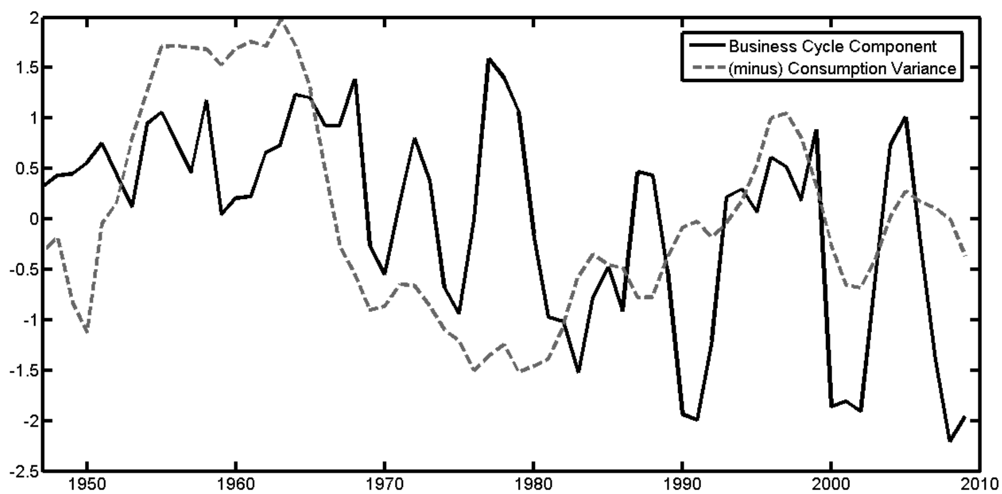

Figure 6: Business cycle component of the short-rate and consumption variance approximated to level of persistence $j = 4$. Both series are demeaned and divided by their standard deviation. The sample spans the period 1947-2010.

Panel A: h-year rolling window

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
v_t^2		1.64	0.56	0.89	0.48	1.35	2.78	2.81	3.75	4.75
	(-)	(1.07)	(0.49)	(0.74)	(0.29)	(0.89)	(2.13)	(1.98)	(2.92)	(4.24)
	{-}	{0.14}	{0.06}	{0.12}	{0.07}	{0.21}	{0.47}	{0.45}	{0.60*}	{0.82**}
$R^2(\%)$	[-0.00]	[2.03]	[0.42]	[1.58]	[0.52]	[4.50]	[18.56]	[17.04]	[27.32]	[41.19]

Panel B: sum of absolute residuals from an AR(1) model of consumption growth

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
v_t^2	-0.11	2.77	2.11	1.12	0.50	1.92	3.96	4.46	5.52	6.49
	(-0.05)	(2.86)	(1.50)	(0.73)	(0.25)	(1.02)	(2.70)	(2.84)	(4.30)	(5.94)
	{-0.01}	{0.26}	{0.23}	{0.13}	{0.06}	{0.24}	{0.55*}	{0.61*}	{0.80**}	{1.04**}
$R^2(\%)$	[0.01]	[6.39]	[5.05]	[1.66]	[0.34]	[5.61]	[23.56]	[28.02]	[39.77]	[52.73]

Panel C: heterogeneous ARCH model

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
v_t^2	1.29	1.40	0.58	0.40	0.35	1.29	2.32	2.61	3.43	4.22
	(1.05)	(1.63)	(0.72)	(0.44)	(0.31)	(1.21)	(2.55)	(2.60)	(3.63)	(4.91)
	{0.12}	{0.19}	{0.10}	{0.08}	{0.07}	{0.26}	{0.52*}	{0.58*}	{0.80**}	{1.08**}
$R^2(\%)$	[1.52]	[3.47]	[0.98]	[0.58]	[0.47]	[6.65]	[21.92]	[25.48]	[39.57]	[54.87]

Table 1: We run linear regressions (with an intercept) of h-period continuously compounded market returns on the CRSP value-weighted index in excess of a 1-year Treasury bill rate on h-period past consumption variance $v_{t-h,t}$. We consider values of h equal to $1-10$ years. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, the $t/\sqrt{T}$ test suggested in Valkanov (2001) in curly brackets, and R^2 statistics in square brackets. Significance at the 5% and 2.5% level of the $t/\sqrt{T}$ test using Valkanovs (2001) critical values is indicated by * and **, respectively. In panel A we use an h-year rolling windows as the proxy for conditional consumption variance v_t^2. In panel B we instead measure consumption volatility $v_{t-h,t} \equiv \log \sum_{i=0}^{h-1} (|\eta_{t-i}|)$, where consumption residuals, η_{t-i} are obtained from $g_t = \mu + \beta g_{t-1} + \eta_t$ See also Bansal, Khatchatrian and Yaron (2005). In panel C the consumption volatility measure is estimated by an HARCH(5) process. The sample is annual and spans the period 1930-2010.

Percentiles	1%	5%	10%	median	90%	95%	99%
Case 1: $\rho = -0.2$							
$R^2(\%)$	0.0	0.0	0.2	5.5	27.9	37.8	53.9
Case 2: $\rho = 0.0$							
$R^2(\%)$	0.0	0.0	0.2	4.7	25.1	33.9	51.5

Table 2: The table reports the percentiles for R^2 in regression (11). The percentiles are calculated based on formula (12) using 100,000 simulations. Integrals are calculated using 1,000 steps per unit interval. It is assumed that the sample consists of 81 yearly observations, and the forecasting horizon is 10 years.

Time scale	Frequency resolution
$j = 1$	$1 - 2$ years
$j = 2$	$2 - 4$ years
$j = 3$	$4 - 8$ years
$j = 4$	$8 - 16$ years
$j = 5$	$16 - 32$ years
π_t^5	> 32 years

Table 3: Interpretation in terms of cycle's duration of the persistence level (or time scales) j in the case of annual time series.

	Level of persistence j				
$z_t =$	1	2	3	4	5
	-2.18	1.06	0.27	0.87	1.09
$v_{j,t}^2$	(-2.34)	(-1.47)	(0.55)	(2.24)	(1.94)
	[0.05]	[0.03]	[0.00]	[0.22]	[0.36]

Table 4: This table reports the results of componentwise predictive regressions of the components of excess stock market returns on the components of consumption variance $v_{j,t}^2$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2010.

Regression: $r_{j,t+2^j}^{(j)} = \alpha + \gamma\sigma_{a,j,t}^2 \; j = 1, \dots, J$		
Sample	$\hat{\alpha}$	$\hat{\gamma}$
1928M1-2010M12	-0.10	2.84
	(-0.12)	(3.91)
1969M1-2010M12	-0.08	2.52
	(0.14)	(2.96)
1928M1-1968M12	0.63	2.68
	(1.85)	(2.26)

Table 5: This table displays the risk aversion estimates based on the persistence heterogeneity tests of the risk-return trade-off. The estimate are based on not-redundant Haar decomposition as suggested in Fadili and Bullmore (2002) based on 512 data points for the first row and 256 data points for the second and third rows.

	Scale j				
$z_t =$	1	2	3	4	5
	-0.00	-0.01	-0.01	-0.01	-0.01
pd_t	(-0.90)	(-5.56)	(-5.23)	(-6.45)	(-2.94)
	[0.00]	[0.16]	[0.20]	[0.07]	[0.10]

Table 6: This table reports the results of regressions of the components of (log) price-dividend ratio $pd_{j,t}$ on the components of consumption growth variance $v^2_{c,j,t}$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2010.

Regression:	$r_{f,j,t+1} = \beta_0 + \beta_1 v^2_{j,t} \;\; j = 1, \ldots, J$	
Sample	$\hat{\beta}_1$	R^2
1930-2010	-1.42	[0.71]
	(-3.72)	
1947-2010	-1.83	[0.65]
	(-4.77)	

Table 7: This table reports the result of componentwise regression of the real short-rate (in percentage) on the components of consumption volatility $v^2_{j,t}$, where the loading coefficient β_1 is restricted to be the same across all levels of persistence. The table reports OLS estimates of the regressors, t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2010.

Panel A: h-year rolling window

	Horizon h (in years)					
	1	2	3	5	7	10
v_t^2		0.86	3.50	6.66	8.76	9.54
	(-)	(0.42)	(1.94)	(4.78)	(5.35)	(7.27)
	{-}	{0.05}	{0.25}	{0.52**}	{0.78**}	{0.71*}
R^2(%)	[]	[0.26]	[5.92]	[21.53]	[38.82]	[34.33]

Panel B: sum of absolute residuals from an AR(1) model

	Horizon h (in years)					
	1	2	3	5	7	10
v_t^2	-0.79	1.28	4.91	8.64	11.46	12.69
	(-0.38)	(0.43)	(1.82)	(3.85)	(4.77)	(5.71)
	{-0.04}	{0.06}	{0.27}	{0.47*}	{0.65**}	{0.55}
R^2(%)	[0.15]	[0.36]	[7.03]	[18.68]	[30.53]	[23.68]

Panel C: heterogeneous ARCH model

	Horizon h (in years)					
	1	2	3	5	7	10
v_t^2	1.24	2.45	2.57	5.08	6.25	7.55
	(0.71)	(1.36)	(1.38)	(3.10)	(5.45)	(6.48)
	{0.07}	{0.17}	{0.22}	{0.52**}	{0.80**}	{1.12**}
R^2(%)	[0.52]	[3.06]	[4.94]	[22.01]	[39.79]	[56.76]

Table 8: We run linear regressions (with an intercept) of h-period continuously compounded market returns $r_{t,t+h}$ on the CRSP value-weighted index in excess of a h-year constant maturity yield on h-period past consumption variance $v_{t-h,t}$. We consider values of h equal to 1-10 years. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, the $t/\sqrt{T}$ test suggested in Valkanov (2001) in curly brackets, and R^2 statistics in square brackets. Significance at the 5% and 2.5% level of the $t/\sqrt{T}$ test using Valkanovs (2001) critical values is indicated by * and **, respectively. In panel A we use an h-year rolling windows as the proxy for conditional consumption variance v_t. In panel B we instead measure consumption volatility $v_{t-h,t} \equiv \log \sum_{i=0}^{h-1} (|\eta_{t-i}|)$, where consumption residuals, $\eta_{c,t-i}$ are obtained from $g_t = \mu + \beta g_{t-1} + \eta_t$ See also Bansal, Khatchatrian and Yaron (2005). In panel C the consumption volatility measure is estimated by an HARCH(5) process. The sample is annual and spans the period 1930-2010.

A Data

We use data on consumption and asset prices for the time period from 1930 till 2010. We take the view that this sample better represents the overall variation in asset and macro economic data. Importantly, the long span of the data helps in achieving more reliable statistical inference. We work with the data sampled on an annual frequency as they are less prone to errors that arise from seasonalities and other measurement problems highlighted in Wilcox (1992).

The data are constructed as follows. For our measure of aggregate consumption we use aggregate consumption of nondurables and services taken from the national income and product accounts (NIPA) tables available from the Bureau of Economic Analysis. This measure is annual from 1929 to 2010, is in real terms, and is seasonally adjusted. The consumption data for UK and Canada are taken from the Barro and Ursua's (2008) cross-country dataset.[33] Growth rates of consumption is constructed by taking the first difference of the corresponding log series.

Stock market returns are computed using the NYSE/AMEX/NASDAQ value-weighted index as market portfolio.[34] These data are provided by CRSP/WRDS and cover the period from January 1926 to December 2010. To compute annual continuously compounded market returns we aggregate monthly continuously compounded return (including dividends) during the t-th month of the year. Price and dividend series are constructed on the per-share basis as in Campbell and Shiller (1988) and Hansen, Heaton, and Li (2008).

The one-year real interest rate series has been obtained from the Shiller data set.[35] Alternatively I follow Bansal et al. (2009) and I construct the ex-ante annual real risk-free rate as the annualized predicted value from a projection of the ex-post real rate on the current three-month nominal yield and inflation over the previous year. The results are robust with regard to this choice.

Since stock returns are measured more frequently than consumption, I obtain a measure of volatility by starting with daily data. In particular to obtain a measure of variance for the

[33]This new dataset is described in Ursua (2011).

[34]Bandi and Perron (2008) and Sizova (2010) use the NYSE/AMEX value-weighted index with dividends as the market proxy. Our results are robust to this choice of the index.

[35]Downloadable at Robert Shiller's homepage http://www.econ.yale.edu/ shiller/

return on the CRSP-VW index, we use the time-series of daily returns:

$$\sigma^2_{m,t,t+1} = \sum_{j=1}^{n_t} r^2_{m,t+\frac{j}{n_t}}$$

where $\sigma^2_{m,t,t+1}$ is the monthly realized variances of the market return in period t and $r_{m,k}$ is the daily CRSP-VW return where k represents a day and n_t is the number of trading days in month t. h-period variance can then be obtained using the data on past realized market variances:

$$\sigma^2_{m,t-h,t} = \sum_{j=1}^{h} \sigma^2_{m,t-i,t-i+1}$$

B Inferential issues and asymptotic approximation

In this section we use a near-unit root specification to derive accurate asymptotic approximations. We rewrite equations (8) and (9) as

$$r^e_{t+1} = \beta v^2_t + \epsilon_{t+1} \tag{B.1}$$

$$(1 - \left(1 + \frac{c}{T}\right))b(L)v^2_t = w_{t+1} \tag{B.2}$$

where the error term ϵ_{t+1} captures the fact that the realized excess returns r^e_{t+1} are a noisy measure of the true unobserved equity premium $E_t[r_{t+1} - r_{f,t}]$. We define $\varphi_v = 1 + \frac{c}{T}$. The parameter c is a constant measuring deviations from unity that are decreasing in T. Thus the process v^2_t is defined as a nearly integrated process, as in Valkanov (2003). Assume the vector $(\epsilon_{t+1}, w_{t+1})$ is a vector martingale difference sequence with covariance matrix $\Sigma \equiv [\sigma^2_\epsilon, \sigma_{\epsilon,w}; \cdot, \sigma^2_w]$. Note that we allow for processes ϵ_{t+1} and w_{t+1} to be (negatively) correlated.

Consider again the regressions in (11). I am interested in the behavior of the R^2 when $\beta_1 = \beta \neq 0$. Under standard OLS assumptions, it is possible to show that

$$\hat{R}^2_h = \frac{\left(\frac{1}{T-2h}\sum_{t=h}^{T-h}(r^e_{t,t+h} - \overline{r^e_{t,t+h}})(v^2_{t-h,t} - \overline{v^2_{t-h,t}})\right)^2}{\frac{1}{T-2h}\left(\sum_{h}^{T-h}(r^e_{t,t+h} - \overline{r^e_{t,t+h}})^2\right)\frac{1}{T-2h}\left(\sum_{t=h}^{T-h} v^2_{t-h,t} - \overline{v^2_{t-h,t}}\right)} \tag{B.3}$$

is a consistent estimate of R_h^2. The $\hat{R}^2$ in regressions (11) have a non-standard distribution for this model because v_t^2 behaves as a unit-root process for $T \to \infty$. The standard assumption in the literature on overlapping observations[36] is that the portion of the overlap h is a constant fraction of the sample size T. Formally, it is captured by the condition $\lim_{T\to\infty} h/T = \lambda$. Under this assumption I show that $\hat{R}^2$ in regression (11) converges to a nondegenerate random variable.

Proof of Theorem 1. The dynamics of the vector (v_t^2, ϵ_{t+1}) is described by the system

$$I - \left(I + \frac{C}{T}L\right)\left(v_t^2, \sum_{i=1}^{t} \epsilon_i\right)^{\mathsf{T}} = (u_t, \epsilon_t).$$

where I is the identity matrix, $u_t \equiv b(L)^{-1} w_t$ satisfies some technical assumptions (see Phillips, 1987). ϵ_t is a martingale difference and C is a diagonal matrix with the elements c and 0. Consider the transformation

$$\left(\frac{v_{[sT]}^2}{\omega\sqrt{T}}, \frac{\sum_{i=1}^{[sT]} \epsilon_i}{\sqrt{T}\sigma_\epsilon}\right)$$

defined on $s \in [0, 1]$ where sT denotes the closest integer that is less than or equal to $s \times T$ and $\omega^2 = \frac{\sigma_w^2}{b(1)^2}$. As follows from Phillips (1987, Lemma 1, p. 539 and 1988, Lemma 3.1, p. 1026), this vector process converges to $(J(s), W_1(s))$ where $J(s)$ is an Ornstein-Uhlenbeck process,

$$dJ(s) = cJ(s) + dW_2(s)$$

with $J(0) = 0$ and where $dW_1(s)$ and $dW_2(s)$ are two standard Weiner processes with covariance $\sigma_{\epsilon,w}/(\sigma_\epsilon\sigma_w)$ and such that $\frac{\sum_{i=1}^{[sT]} w_i}{\sqrt{T}\sigma_w}$ weakly converges to $W_2(s)$.

Define $t = \tau T$ and $h = \lambda T$ in the original parametrization. The asymptotic distribution of the first (predictable) part of the multi-period return $\beta \sum_{i=1}^{h} v_{t+i-1}^2$ follows from the continuous mapping theorem (CMT),

$$\beta \frac{\sum_{i=[\tau T]}^{[(\tau+\lambda)T]} v_i^2}{\omega T^{3/2}} = \beta \frac{\int_\tau^{\tau+\lambda} v_{[sT]}^2 ds}{\omega\sqrt{T}} \Rightarrow \beta \int_\tau^{\tau+\lambda} J(s)ds \qquad \text{(B.4)}$$

[36] Differently from Valkanov's framework and standard literature on spurious regressions, however, regressor and regressand are aggregated over non-overlapping periods.

which requires normalization by $T^{3/2}$, in contrast to T, which is a normalization coefficient for the unpredictable part. Therefore, as $T \to \infty$ the first (predictable) part dominates the second one, and the limit of the multi-period return is the same as the limit in equation (B.4), i.e.

$$\frac{r^e_{t,t+h}}{\omega T^{3/2}} \Rightarrow \beta \int_{\tau}^{\tau+\lambda} J(s)ds$$

The rest of the steps are also based on the CMT. The demeaned returns converge to the demeaned version of the above integral,

$$\frac{r^e_{t,t+h} - \frac{1}{T-2h}\sum_{t=h}^{T-h}\left(r^e_{t,t+h}\right)}{\omega T^{3/2}} \Rightarrow \beta \underbrace{\int_{\tau}^{\tau+\lambda} J(s)ds - \frac{1}{1-2\lambda}\int_{\lambda}^{1-\lambda}\int_{r}^{r+\lambda} J(s)dsdr}_{A(\tau)} \tag{B.5}$$

The asymptotic behavior of the past multi-period variance follows from (B.4)

$$\frac{\sum_{i=[(\tau-\lambda)T]}^{[\tau T]} v_i^2}{\omega T^{3/2}} = \beta \frac{\int_{\tau-\lambda}^{\tau} v^2_{[sT]}ds}{\omega\sqrt{T}} \Rightarrow \beta \int_{\tau-\lambda}^{\tau} J(s)ds$$

The limit for its demeaned version again follows from the CMT,

$$\frac{v^2_{t-h,t} - \frac{1}{T-2h}\sum_{t=h}^{T-h}\left(v^2_{t-h,t}\right)}{\omega T^{3/2}} \Rightarrow \beta \underbrace{\int_{\tau-\lambda}^{\tau} J(s)ds - \frac{1}{1-2\lambda}\int_{\lambda}^{1-\lambda}\int_{r-\lambda}^{r} J(s)dsdr}_{B(\tau)} \tag{B.6}$$

Finally note that $\hat{R}^2_h$ can be represented as $\hat{R}^2_h = F_{R^2}(A_T, B_T)$ where

$$A_T(\tau) = \frac{r^e_{[\tau T],[(\tau+\lambda)T]} - \frac{1}{T-2h}\sum_{i=[\lambda T]}^{T-[\lambda T]}\left(r^e_{i,i+h}\right)}{\omega T^{3/2}}$$

$$B_T(\tau) = \frac{v^2_{[(\tau-\lambda)T],[\tau T]} - \frac{1}{T-2h}\sum_{i=[\lambda T]}^{T-[\lambda T]}\left(v^2_{i-h,h}\right)}{\omega T^{3/2}}$$

$$F_{R^2}(X,Y) = \frac{\left(\int_{\lambda}^{1-\lambda} X(\tau)Y(\tau)d\tau\right)^2}{\left(\int_{\lambda}^{1-\lambda} X^2(\tau)d\tau\right)\left(\int_{\lambda}^{1-\lambda} Y^2(\tau)d\tau\right)}$$

By applying the CMT once more, we have $\hat{R}^2_h \Rightarrow F_{R^2}(A,B)$. □

Using the expressions for $A(\tau)$ and $B(\tau)$ and formula (12) I construct the distribution for the $\hat{R}^2_h$. The coefficients are chosen as follows. $c = (0.75.1)T$, which corresponds to the observed first autocorrelation of 0.75 annual consumption volatility. The correlation coefficient $\sigma_{\epsilon,w}/(\sigma_\epsilon\sigma_w)$, is fixed at -0.2. For comparison, we also report the case with no leverage.

C Robustness

In this Appendix, we report some of the robustness checks that we have conducted.

First we check whether the use of one-year (real) interest rate compounded over horizon h as a proxy for the risk-free rate at horizon h can possibly drive our results. In fact we would like to possibly use a risk-free rate with maturity equal to that of the h-year holding period returns. We thus repeat the exercise in Table 1 where we use the yield-to-maturity on a real h-year US Treasury bond[37] as the the risk-free rate used to compute the excess returns at horizon h.

[Insert Table 8 about here.]

The general pattern of results using this method is very similar to those using the compounded one-year risk-free rate.

Next we consider a potential econometric hazard with interpreting the long-horizon regression results. In particular the use of overlapping data in long-horizon regressions can skew statistical inference in finite samples. It is possible to show that the standard t-statistic of long-horizon regression coefficients diverges with the sample size T, thereby determining likely over-rejections in the classical asymptotic framework. However we can address this potential inference problem by noting that the rescaled $t/\sqrt{T}$ statistic has a well-defined limiting distribution and can be used to test the null of no dependence. This is similar to Valkanov (2003) although we use a different aggregation method where regressor and regressand are aggregated over non-overlapping periods.

The distribution of this rescaled statistic is nonstandard, however, and depends on two nuisance parameters, δ and c. The parameter δ measures the covariance between innovations in the variable to be forecast, and innovations in some forecasting variable, call it X_t. The parameter c measures deviations from unity in the highest autoregressive root for X_t, in a decreasing (at rate T) neighborhood of 1. We assume the parameters are known: $c = (\rho_1 - 1)T$ with $\rho_1 = 0.75$ and $delta = -0.2$. In the relevant region of the parameter space, however, we find that the distribution is not very sensitive to these values. With these

[37]real yield = (nominal yield) - (inflation rate).

parameters in hand, the rescaled t-statistic, $t/\sqrt{T}$, I can generate critical values. Tables 1 and 8 contains inference based on the $t/\sqrt{T}$ statistic. The rescaled t-statistics for our application are reported in curly brackets. The tables report both the statistic itself and whether its value implies that the predictive coefficient in each regression is statistically significant at the 5% and 2.5% levels. The dependence between excess returns and past variance is still very significant at the 5% level over 7, 8, 9, and 10 years. According to these statistics, the forecasting power of past consumption volatility for long-horizon excess stock market returns is robust to accounting for biases arising from the use of overlapping data in finite samples.

Finally we re-run the regressions in Tables 1 and 8 using variance instead of volatility. This matters little to our results.

D The Valuation Approach: the details of the derivation

In this Section we show the steps to obtain the values of the financial ratios coefficients $A_{0,j}, A_j, A^m_{0,j}, A^m_j$ in terms of the parameters of the model. We then compute the equity premia on both the consumption claim asset and the market return. Finally we derive the risk-free rate. In what follows, we make use of the decomposition of time series into layers with different levels of persistence described in Section 4 and reported here for the reader's convenience:

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_t^J \tag{D.1}$$

Alternatively, to reconstruct the realization of x_{t+1} from the effect that this realization will have at different horizons we use

$$x_{t+1} = \sum_{j=1}^{J} x_{t+2^j}^{(j)} + \pi_{t+2^J}^{(J)} \tag{D.2}$$

D.1 The Financial Ratios

We solve first for the price-consumption coefficients $A_{0,j}, A_j$ and hence for the consumption return $r_{a,t+1}$. This determines the pricing kernel.

To obtain the values of the coefficients $A_{0,j}, A_j$ we exploit the Euler condition

$$
\begin{aligned}
&E_t\left[\exp\left(\theta\log\beta-\frac{\theta}{\psi}g_{t+1}+(\theta-1)r_{a,t+1}+r_{i,t+1}\right)\right]\\
=&E_t\left[\exp\left(\theta\log\beta-\frac{\theta}{\psi}g_{t+1}+\theta r_{a,t+1}\right)\right]=1
\end{aligned}
$$

which is derived from (25) for the special case where the asset being priced is the aggregate consumption claim, i.e. $r_{i,t+1}=r_{a,t+1}$. We then express the log consumption growth g_{t+1} and the return $r_{a,t+1}$ in terms of the factors $\{x_{j,t}\}_j$ and of the innovations $\{e^g_{j,t+2^j}\}_j$ and $\{\varepsilon_{j,t+2^j}\}_j$. To do so we plug first the Campbell and Shiller (1988) approximation for log returns, see equation (27), into the above expression to obtain:

$$
E_t\left[\exp\left(\theta\log\beta-\frac{\theta}{\psi}g_{t+1}+\theta\underbrace{\left(\kappa_0+\kappa_1 z^a_{t+1}-z^a_t+g_{t+1}\right)}_{r_{a,t+1}}\right)\right]=1
$$

By the backward decomposition (D.1) applied to the (demeaned) price-consumption ratio at time t and by the forward decomposition (D.2) applied to the (demeaned) consumption growth and price-consumption processes at time $t+1$ we have:

$$
z_{a,t}=\sum_{j=1}^{J} z^{(j)}_{a,t} \tag{D.3}
$$

$$
z_{a,t+1}=\sum_{j=1}^{J} z^{(j)}_{a,t+2^j} \tag{D.4}
$$

$$
g_{t+1}=\sum_{j=1}^{J} g^{(j)}_{t+2^j} \tag{D.5}
$$

We now show that the variance of the consumption component at level j, $\sigma^2_{j,t}$, coincides with the component at level j of the consumption variance. In fact we have that:

$$
\sigma^2_t=Var_t(g_{t+1})=Var_t\left(\sum_{j=1}^{J} g^{(j)}_{t+2^j}\right)=Var_t\left(\sum_{j=1}^{J}\sigma_{j,t}\,e^g_{j,t+2^j}\right)=\sum_{j=1}^{J}\sigma^2_{j,t}
$$

where we use the fact that the shocks $e_{j,t}$ and $e_{j',t}$ are uncorrelated for all $j\neq j'$.

Plugging the above expressions into the Euler condition yields:

$$
E_t\left[\exp\left(\theta\log\beta-\frac{\theta}{\psi}\left(\sum_{j=1}^{J} g^{(j)}_{t+2^j}\right)+\theta\left(\kappa_0+\kappa_1\left(\sum_{j=1}^{J} z^{(j)}_{t+2^j}\right)-\left(\sum_{j=1}^{J} z^{(j)}_{t}\right)+\left(\sum_{j=1}^{J} g^{(j)}_{t+2^j}\right)\right)\right)\right]=1
$$

Finally using the dynamics for the components of log consumption growth given in equation (22) together with our guess for the components of price-consumption ratio solution given in equation (28), rearranging terms and using the log normal properties of the shocks we obtain:

$$E_t\left[\exp\left(\theta\log\beta+\theta\left(1-\frac{1}{\psi}\right)\left(\sum_{j=1}^{J}g^{(j)}_{t+2^j}\right)+\theta\left(\kappa_0+\kappa_1\left(\sum_{j=1}^{J}z^{(j)}_{t+2^j}\right)-\left(\sum_{j=1}^{J}z^{(j)}_t\right)\right)\right)\right]$$

$$=E_t\Bigg[\exp\quad\Bigg(\theta\log\beta+\theta\left(1-\frac{1}{\psi}\right)\left(\sum_{j=1}^{J}\sigma_{j,t}e^g_{j,t+2^j}\right)+$$

$$\theta\left(\kappa_0+\kappa_1\left(\sum_{j=1}^{J}A_{0,j}+\sum_{j=1}^{J}A_j\sigma^2_{j,t+2^j}\right)-\left(\sum_{j=1}^{J}A_{0,j}+\sum_{j=1}^{J}A_j\sigma^2_{j,t}\right)\quad\right)\quad\Bigg]$$

$$=E_t\Bigg[\exp\quad\Bigg(\quad\theta(\log\beta+\kappa_0+(\kappa_1-1)\sum_{j=1}^{J}A_{0,j})+\dots$$

$$\theta\left(1-\frac{1}{\psi}\right)\left(\sum_{j=1}^{J}\sigma_{j,t}e^g_{j,t+2^j}\right)+\theta\left(\kappa_1\sum_{j=1}^{J}A_j\sigma^2_{j,t+2^j}-\sum_{j=1}^{J}A_j\sigma^2_{j,t}\right)\quad\Bigg)\quad\Bigg]$$

$$=E_t\Bigg[\exp\quad\Bigg(\quad\theta(\log\beta+\kappa_0+(\kappa_1-1)\sum_{j=1}^{J}A_{0,j})+\dots$$

$$\theta\left(1-\frac{1}{\psi}\right)\left(\sum_{j=1}^{J}\sigma_{j,t}e^g_{j,t+2^j}\right)+\theta\left(\kappa_1\sum_{j=1}^{J}A_j\underbrace{(e_jM\tilde{\Sigma}_t+e_j\varepsilon_{t+2^J})}_{\sigma^2_{j,t+2^j}}-\sum_{j=1}^{J}A_j\sigma^2_{j,t}\right)\quad\Bigg)\quad\Bigg]=1$$

where we defined $\tilde{\Sigma}_t\equiv[\sigma^2_{1,t},\dots,\sigma^2_{J,t}]^\intercal$. Collecting terms in $\tilde{\Sigma}_t$ yields eventually a system of equations

$$e_j\left(0.5\left(\theta-\frac{\theta}{\psi}\right)^2+\theta A_j(\kappa_1M-\mathbb{I}_J)\right)=0$$

for all $j=1,\dots,J$. If we introduce the following column vectors

$$\underline{A}\ \equiv\ [A_1,\dots,A_J]^\intercal$$

the solution to these equations is given by the following vectors of sensitivities:

$$\underline{A}\ =\ 0.5\frac{\left(\theta-\frac{\theta}{\psi}\right)^2}{\theta}\left(\mathbb{I}_J-\kappa_1M\right)^{-1}\underline{1}$$

D.2 The Risk Premium and Return Volatility

The risk premium for any asset is determined by the conditional covariance between the return and the SDF. For instance we can compute the risk premium on any asset i as

$$E_t[r_{i,t+1} - r_{f,t}] + 0.5\sigma^2_{r_{i,t}} = -cov_t(m_{t+1}, r_{i,t+1})$$

We therefore need to compute first the innovations in the stochastic discount factor and in the returns.

Given the solution above for $z^{(j)}_{a,t}$ it is possible to derive the innovation to the return $r_{a,t+1}$ as a function of the evolution of the state variables and the parameters of the model. In particular the equilibrium return innovations can be found by plugging the expressions (D.3), (D.4) and (D.5) into the Campbell and Shiller (1988) approximation for log returns, see equation (27) to obtain

$$\begin{aligned} r_{a,t+1} - E_t[r_{a,t+1}] &= \left(\sum_{j=1}^{J} g^{(j)}_{t+2^j}\right) + \kappa_0 + \kappa_1 \left(\sum_{j=1}^{J} z^{(j)}_{t+2^j}\right) - \left(\sum_{j=1}^{J} z^{(j)}_{t}\right) - E_t[r_{a,t+1}] \\ &= \sum_{j=1}^{J} \sigma_{j,t} e^g_{j,t+2^j} + \kappa_1 \left(\sum_{j=1}^{J} A_j (e_j \varepsilon_{t+2^J})\right) \\ &= \sigma_{j,t} \odot e^g_{j,t+1} + \kappa_1 \underline{A} \boldsymbol{\varepsilon}_{t+1} \end{aligned} \tag{D.6}$$

where we define

$$\begin{aligned} \boldsymbol{\varepsilon}^\intercal_{t+1} &\equiv [\varepsilon_{1,t+2^1}, \ldots, \varepsilon_{J,t+2^J}] \\ \sigma_{j,t} \odot e^g_{j,t+1} &\equiv \sum_{j=1}^{J} \sigma_{j,t} e^g_{j,t+2^j} \end{aligned}$$

The innovation in the return component at level of persistence j is given by:

$$r^{(j)}_{a,t+2^j} - E_t[r^{(j)}_{a,t+2^j}] = \sigma_{j,t} e^g_{j,t+2^j} + \kappa_1 \left(A_j \varepsilon_{j,t+2^j}\right) \tag{D.7}$$

It is trivial to show that

$$r_{a,t+1} - E_t[r_{a,t+1}] = \sum_{j=1}^{J} r^{(j)}_{a,t+2^j} - E_t[r^{(j)}_{a,t+2^j}]$$

which allows us to decompose the innovation in aggregate return into the sum of the innovation in the market return components. Further, it follows that the conditional variance of $r_{a,t+1}$ is:

$$Var_t(r_{a,t+1}) = \sum_{j=1}^{J} \sigma_{j,t}^2 + \kappa_1^2 \underline{A}\mathbf{Q}\underline{A}' \tag{D.8}$$

where we define

$$\mathbf{Q} \equiv \mathbf{E}_t \left[\boldsymbol{\varepsilon}_{t+1} \boldsymbol{\varepsilon}'_{t+1} \right]$$

To find the innovations in the stochastic discount factor, we plug the expressions (D.3), (D.4) and (D.5), together with the dynamics for the components of log consumption growth given in equation (22) and our guess for the components of price-consumption ratio solution given in equation (28) into equation (26) to obtain:

$$\begin{aligned}
m_{t+1} &= \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} \\
&= \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1)(\kappa_0 + \kappa_1 z_{t+1}^a - z_t^a + g_{t+1}) \\
&= \theta \log \beta - \frac{\theta}{\psi} \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (\theta - 1) \left(\kappa_0 + \kappa_1 \sum_{j=1}^{J} z_{t+2^j}^{(j)} - \sum_{j=1}^{J} z_t^{(j)} + \sum_{j=1}^{J} g_{t+2^j}^{(j)} \right) \\
&= \theta log(\beta) - \left(1 - \theta + \frac{\theta}{\psi}\right) \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (\theta - 1) \left(\kappa_0 + \kappa_1 \sum_{j=1}^{J} z_{t+2^j}^{(j)} - \sum_{j=1}^{J} z_t^{(j)} \right) \\
&= \theta log(\beta) - \left(1 - \theta + \frac{\theta}{\psi}\right) \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (\theta - 1) \left(\kappa_0 + \kappa_1 \sum_{j=1}^{J} z_{t+2^j}^{(j)} - \sum_{j=1}^{J} z_t^{(j)} \right) \\
&= \theta log(\beta) - \left(1 - \theta + \frac{\theta}{\psi}\right) \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (\theta - 1) \left(\kappa_0 + \kappa_1 \sum_{j=1}^{J} A_{0,j} + A_j \sigma_{j,t+2^j}^2 - \sum_{j=1}^{J} A_{0,j} - A_j \sigma_{j,t}^2 \right)
\end{aligned}$$

Finally using the dynamics for our latent factors (23) we obtain

$$\begin{aligned}
m_{t+1} = \theta log(\beta) + (\theta - 1) &\left(\kappa_0 + \kappa_1 \sum_{j=1}^{J} A_{0,j} + A_j \rho_j \sigma_{j,t}^2 - \sum_{j=1}^{J} A_{0,j} - A_j \sigma_{j,t}^2 \right) \\
&- \left(1 - \theta + \frac{\theta}{\psi}\right) \sum_{j=1}^{J} \sigma_{j,t} e_{j,t+2^j}^g + (\theta - 1)\kappa_1 \left(\sum_{j=1}^{J} A_j \varepsilon_{j,t+2^j} \right)
\end{aligned}$$

which implies

$$
\begin{aligned}
m_{t+1} - E_t[m_{t+1}] &= -\left(1-\theta+\frac{\theta}{\psi}\right)\sum_{j=1}^{J}\sigma_{j,t}e^{g}_{j,t+2^j} + (\theta-1)\kappa_1\left(\sum_{j=1}^{J}A_j\varepsilon_{j,t+2^j}\right) \\
&= -\lambda_g\sum_{j=1}^{J}\sigma_{j,t}e^{g}_{j,t+2^j} - \sum_{j=1}^{J}\lambda_j\varepsilon_{j,t+2^j} \\
&= -\lambda_g\sigma_{j,t}\odot e^{g}_{j,t+1} - \underline{\lambda}_{\varepsilon}\boldsymbol{\varepsilon}_{t+1}
\end{aligned}
\tag{D.9}
$$

where

$$
\begin{aligned}
\lambda_g &\equiv \left(\frac{\theta}{\psi}-\theta+1\right) = \gamma \\
\underline{\lambda}_{\varepsilon} &\equiv \kappa_1(1-\theta)\underline{A}
\end{aligned}
$$

Using the formula (D.7) and the innovation in the SDF (D.9) we obtain the risk premium for the components of consumption claim asset

$$
E_t[r_{a,t+2^j} - r_{f,t+2^j}] + 0.5\sigma^2_{r_{a,t,j}} = \lambda_g\sigma^2_{j,t} + \kappa_1[\underline{\lambda}_{\varepsilon}\mathbf{Q}]_j A_j
$$

and using the formula for the return on aggregate wealth (D.6) and the innovation in the SDF (D.9) we obtain the risk premium for the consumption claim asset,

$$
E_t[r_{a,t+1} - r_{f,t}] + 0.5\sigma^2_{r_{a,t}} = \lambda_g\sum_{j=1}^{J}\sigma^2_{j,t} + \kappa_1\underline{\lambda}_{\varepsilon}\mathbf{Q}\underline{A}'
$$

where $\sigma^2_{r_{a,t}}$ is defined in equation (D.8).

D.3 Risk-Free Rate Dynamics

To obtain our expression for the risk-free rate we start by plugging the log short-term real interest rate $r_{f,t+1}$ for r^i_{t+1} into the Euler equation (25). Then by applying the forward decomposition (D.2) to the (demeaned) consumption growth and to the log returns processes at time $t+1$ we observe that the risk-free rate between t and $t+1$, $r_{f,t+1}$ satisfies the following condition:

$$
E_t\left[\exp\left(\theta\log\beta - \left(\frac{\theta}{\psi}\right)\sum_{j=1}^{J}g^{(j)}_{t+2^j} + (\theta-1)\sum_{j=1}^{J}r_{a,j,t+2^j}\right)\right] = \exp(-r_{f,t+1})
$$

where once again $r_{a,t+1}$ is the return on the asset that pays consumption as dividend. Taking logs on both sides and using the log normal properties of the shocks we can rewrite it as follows

$$\begin{aligned} r_{f,t+1} &= -\theta \log \beta + \frac{\theta}{\psi} E_t \left[\sum_{j=1}^{J} g_{t+2^j}^{(j)} \right] + (1-\theta) E_t \left[\sum_{j=1}^{J} r_{a,j,t+2^j} \right] \\ &- \frac{1}{2} var_t \left[\frac{\theta}{\psi} \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (1-\theta) \sum_{j=1}^{J} r_{a,j,t+2^j} \right] \\ &= -\log \beta + \frac{1}{\psi} E_t \left[\sum_{j=1}^{J} g_{t+2^j}^{(j)} \right] + \frac{(1-\theta)}{\theta} E_t \left[\sum_{j=1}^{J} r_{a,j,t+2^j} - r_f \right] \\ &- \frac{1}{2\theta} var_t \left[\frac{\theta}{\psi} \sum_{j=1}^{J} g_{t+2^j}^{(j)} + (1-\theta) \sum_{j=1}^{J} r_{a,j,t+2^j} \right] \end{aligned} \tag{D.10}$$

where in the last line we subtract $(1-\theta) r_{f,t}$ from both sides and divide by θ, where it is assumed that $\theta \neq 0$. Further to solve the above expression, note that

$$var_t \left[\frac{\theta}{\psi} \sum_{j=1}^{h} g_{j,t+h} + (1-\theta) \sum_{j=1}^{h} r_{a,j,t+h} \right] = var_t(m_{t+1})$$

Recall from (D.9) that

$$m_{t+1} - E_t[m_{t+1}] = -\lambda_g \sigma_{j,t} \odot e_{j,t+1}^{g} - \underline{\lambda}_{\varepsilon} \boldsymbol{\varepsilon}_{t+1}$$

and therefore

$$var_t(m_{t+1}) = \lambda_g^2 \sum_{j=1}^{J} \sigma_{j,t}^2 + \kappa_1^2 (1-\theta)^2 \underline{A} \mathbf{Q} \underline{A}'$$

Note that given the dynamics for log consumption growth (see equation (22)) we have that $E_t \left[\sum_{j=1}^{J} g_{t+2^j}^{(j)} \right] = 0$. Eventually, using the expression for the equity premium of return on aggregate wealth we obtain:

$$r_{f,t+1} = -\log \beta + \lambda_g^2 \sum_{j=1}^{J} \sigma_{j,t}^2 + \kappa_1^2 (1-\theta)^2 \underline{A} \mathbf{Q} \underline{A}' + \frac{(1-\theta)}{\theta} \left(\lambda_g \sum_{j=1}^{J} \sigma_{j,t}^2 + \kappa_1 \underline{\lambda}_{\varepsilon} \mathbf{Q} \underline{A}' \right)$$

References

Abel, Andrew B., "Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model," *Journal of Monetary Economics*, 1988, *22* (3), 375–393.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold, "Parametric and Nonparametric Volatility Measurement," *Handbook of Financial Econometrics*, 2003.

___ , ___ , ___ , and Paul Labys, "Modeling and Forecasting Realized Volatility," *Econometrica*, 2003, *71* (2), pp. 579–625.

Atkeson, Andrew and Patrick J. Kehoe, "On the Need for a New Approach to Analyzing Monetary Policy," NBER Working Papers 14260, National Bureau of Economic Research, Inc August 2008.

Baillie, Richard T. and Ramon P. DeGennaro, "Stock Returns and Volatility," *Journal of Financial and Quantitative Analysis*, June 1990, *25* (02), 203–214.

Bandi, Federico M. and Benot Perron, "Long-run risk-return trade-offs," *Journal of Econometrics*, April 2008, *143* (2), 349–374.

___ , Rene Garcia, Abraham Lioui, and Benoit Perron, "A Long-Horizon Perspective on the Cross-Section of Expected Returns," *SSRN eLibrary*, 2010.

Bansal, Ravi, Amir Yaron, and Dana Kiku, "Risk for the Long Run: Estimation and Inference," Working Paper September 2009.

___ and ___ , "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 08 2004, *59* (4), 1481–1509.

___ , Varoujan Khatchatrian, and Amir Yaron, "Interpretable asset markets?," *European Economic Review*, 2005, *49* (3), 531 – 560.

Barndorff-Nielsen, Ole E. and Neil Shephard, "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics," *Journal Of The Royal Statistical Society Series B*, 2001, *63* (2), 167–241.

Barsky, Robert B., "Why Don't the Prices of Stocks and Bonds Move Together?," *The American Economic Review*, 1989, *79* (5), pp. 1132–1145.

Benigno, Pierpaolo, Luca Antonio Ricci, and Paolo Surico, "Unemployment and Productivity in the Long Run: The Role of Macroeconomic Volatility," NBER Working Papers 16374, National Bureau of Economic Research, Inc September 2010.

Bollerslev, Tim, George Tauchen, and Hao Zhou, "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, November 2009, *22* (11), 4463–4492.

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, "Nonlinearities in the relation between the equity risk premium and the term structure," *Manage. Sci.*, March 1997, *43*, 371–385.

Calvet, Laurent E. and Adlai J. Fisher, "Multifrequency news and stock returns," *Journal of Financial Economics*, October 2007, *86* (1), 178–212.

Campbell, John Y, "Intertemporal Asset Pricing without Consumption Data," *American Economic Review*, June 1993, *83* (3), 487–512.

Campbell, John Y. and Ludger Hentschel, "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, June 1992, *31* (3), 281–318.

Campello, Murillo, Long Chen, and Lu Zhang, "Expected returns, yield spreads, and asset pricing tests," *Review of Financial Studies*, May 2008, *21* (3), 1297–1338.

Canzoneri, Matthew B., Robert E. Cumby, and Behzad T. Diba, "Euler equations and money market interest rates: A challenge for monetary policy models," *Journal of Monetary Economics*, October 2007, *54* (7), 1863–1881.

Chernov, Mikhail, Ronald Gallant, Eric Ghysels, and George Tauchen, "Alternative models for stock price dynamics," *Journal of Econometrics*, 2003, *116* (1-2), 225–257.

Corradi, Valentina, Antonio Mele, and Walter Distaso, "Macroeconomic Determinants of Stock Market Returns, Volatility and Volatility Risk-Premia," FMG Discussion Papers dp616, Financial Markets Group June 2008.

Corsi, Fulvio, "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 2009, *7* (2), 174–196.

Dijkerman, R.W. and R.R. Mazumdar, "Wavelet representations of stochastic processes and multiresolution stochastic models," *Signal Processing, IEEE Transactions on*, jul 1994, *42* (7), 1640 –1652.

Drechsler, Itamar and Amir Yaron, "What's Vol Got to Do with It," *Review of Financial Studies*, 2011, *24* (1), 1–45.

Engle, Robert F. and Gary G. Lee, "A Permanent and Transitory Component Model of Stock Return Volatility," *SSRN eLibrary*, 1993.

Epstein, Larry G and Stanley E Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, July 1989, *57* (4), 937–69.

___ **and** ___ , "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, April 1991, *99* (2), 263–86.

Fadili, M.J and E.T Bullmore, "Wavelet-Generalized Least Squares: A New BLU Estimator of Linear Regression Models with 1/f Errors," *NeuroImage, Volume 15, Number 1*, 2002.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, "Expected stock returns and volatility," *Journal of Financial Economics*, September 1987, *19* (1), 3–29.

Ghysels, Eric, Arthur Sinko, and Rossen Valkanov, "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 2007, *26* (1), 53–90.

___ **, Pedro Santa-Clara, and Rossen Valkanov**, "There is a Risk-Return Tradeoff After All," *Journal of Financial Economics*, 2005, *76* (3), 509 – 548.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle, "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *The Journal of Finance*, 1993, *48* (5), pp. 1779–1801.

Hall, Robert E, "Intertemporal Substitution in Consumption," *Journal of Political Economy*, April 1988, *96* (2), 339–57.

Hansen, Lars Peter, "Modeling the Long Run: Valuation in Dynamic Stochastic Economies," NBER Working Papers 14243, National Bureau of Economic Research, Inc August 2008.

Harvey, Campbell R., "The specification of conditional expectations," *Journal of Empirical Finance*, December 2001, *8* (5), 573–637.

Kandel, Shmuel and Robert F. Stambaugh, "Expectations and Volatility of Long-Horizon Stock Returns," *The Review of Financial Studies*, 1990, (3), 207 – 232.

Lettau, Martin and Sydney C. Ludvigson, "Measuring and Modelling Variation in the Risk-Return Trade-off," *Handbook of Financial Econometrics*, 2010, *1*, 617–690.

——, ——, and Jessica A. Wachter, "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," *Review of Financial Studies*, July 2008, *21* (4), 1653–1687.

Levan, Nhan and Carlos S. Kubrusly, "Multiresolution approximation scale and time-shift subspaces," *Multidimensional Systems and Signal Processing*, April 2006, *17* (4), 343–354.

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

Maheu, John M. and Thomas H. McCurdy, "Components of Market Risk and Return," *Journal of Financial Econometrics*, Fall 2007, *5* (4), 560–590.

Malkhozov, Aytek and Maral Shamloo, "Asset Prices in a News Driven Real Business Cycle Model," Technical Report March 2010.

Mallat, S. G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, July 1989, *11*, 674–693.

Muller, Ulrich A., Michel M. Dacorogna, Rakhal D. Dave, Richard B. Olsen, Olivier V. Pictet, and Jacob E. von Weizsacker, "Volatilities of different time resolutions – Analyzing the dynamics of market components," *Journal of Empirical Finance*, June 1997, *4* (2-3), 213–239.

Ortu, Fulvio, Andrea Tamoni, and Claudio Tebaldi, "Long-Run and the Temporal Aggregation of Risks, `http://www.igier.unibocconi.it/finance`," Technical Report March 2011.

——, ——, and ——, "Long-Run Risk and the Persistence of Consumption Shocks," Working Papers 14260 August 2011.

Parker, Jonathan A. and Christian Julliard, "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, February 2005, *113* (1), 185–222.

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

Schwert, G William, "Why Does Stock Market Volatility Change over Time?," *Journal of Finance*, December 1989, *44* (5), 1115–53.

Sizova, Natalia, "Long-Horizon Return Regressions with Historical Volatility and Other Long-Memory Variables," *SSRN eLibrary*, 2010.

Valkanov, Rossen, "Long Horizon Regressioms: Theoretical Results and Applications," *Journal of Financial Economics*, 2003, *68*, 201–232.

Whitelaw, Robert F, "Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns," *Journal of Finance*, June 1994, *49* (2), 515–41.

Zhou, Guofu and Yingzi Zhu, "A Long-run Risks Model with Long- and Short-Run Volatilities: Explaining Predictability and Volatility Risk Premium," Working Paper September 2010.

Zhou, Hao, "Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty," *SSRN eLibrary*, 2010.