

# Do consumption-based asset pricing models explain serial dependence in stock returns?\*

Michael Ashby<sup>†</sup>      Oliver Linton<sup>‡</sup>

First version: October 2017

This version: May 2018

## Abstract

We show that the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models of asset prices cannot explain the serial correlation structure of stock returns. We show this by estimating these models and deriving expected returns from them and then testing whether the difference between observed and expected returns is a martingale difference sequence. We use variance ratio and rescaled range tests which we modify to account for the expected returns being functions of estimated parameters. We also use a weighted quantilogram test based on a bootstrap procedure robust to this estimation. The evidence against the Bansal-Yaron and Campbell-Cochrane models is significant. While the evidence against the Cecchetti-Lam-Mark model is not in general significant, our point estimates strongly suggest its residuals are not a martingale difference sequence. Furthermore, a semi-parametric maximal predictability test suggests there is some evidence that the three models' state variables struggle to explain the degree of predictability observed in the market return. A timing strategy designed to exploit predictability in the market can significantly outperform the market in certainty equivalent terms under the Bansal-Yaron model. The timing strategy may underperform the market by less than it ought to under the Campbell-Cochrane model.

**JEL classification:** C58, G12

**Keywords:** consumption-based asset pricing models, serial correlation, predictability, martingale difference sequence, variance ratio, quantilogram, rescaled range, mean reversion

## 1 Introduction

It is an apparent stylised fact of financial markets that stock price movements are approximately serially uncorrelated at short horizons, but negatively serially correlated at

---

\*We thank Tiago Cavalcanti, Donald Robertson, Sonje Reiche, Melvyn Weeks and especially Mark Salmon for helpful comments and feedback which have greatly improved this paper, as well as participants at the Cambridge Econometrics Workshop. Ashby gratefully acknowledges financial support from the Economic and Social Research Council under award 1515004.

<sup>†</sup>Faculty of Economics, University of Cambridge. Corresponding author. Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, UK. Email [mwa22@cam.ac.uk](mailto:mwa22@cam.ac.uk).

<sup>‡</sup>Faculty of Economics, University of Cambridge. Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, UK. Email [ob120@cam.ac.uk](mailto:ob120@cam.ac.uk).

longer horizons (Campbell, 2014; RSAS, 2013). However, relatively little work explores whether such a serial correlation structure is compatible with popular models of asset pricing. In this article, we show that three popular asset pricing models are unable to explain the serial correlation structure of observed stock returns. We find that the difference between expected returns implied by the models considered and observed returns is forecastable over time with respect to its own past history. This difference is significantly forecastable for the Bansal-Yaron and Campbell-Cochrane models. While the difference is not in general significantly forecastable for the Cecchetti-Lam-Mark model, our point estimates, coupled with the known low power of the tests used (Poterba and Summers, 1988), strongly suggest the difference is negatively serially correlated. Since the models we test have rational expectations, if the models were correctly specified, this difference would be unforecastable, and so a martingale difference sequence (MDS) at the very least. We estimate the models considered to obtain the model-implied expected returns and modify existing tests of the MDS null to account for this estimation. We also find some evidence that the degree of predictability observed exceeds the maximum possible if the stochastic discount factor (SDF) contains the state variables of the models considered. Moreover, we show that a timing strategy designed to exploit return predictability can yield a certainty equivalent gain under the Bansal-Yaron model, and may reduce the certainty equivalent less than it ought to under the Campbell-Cochrane model.

The extent to which rational models of asset prices can explain stock return predictability of is interest for several reasons. First, most of the models we consider are designed to capture the equity premium (Mehra and Prescott, 1985) and risk-free rate (Weil, 1989) puzzles. Assuming a standard endowment economy with a representative investor who has CRRA<sup>1</sup> preferences, the observed difference between stock returns and low-risk bond yields requires extremely high levels of risk aversion to explain. This is the equity premium puzzle. The risk-free rate puzzle compounds the equity premium puzzle. If CRRA investors are indeed as risk averse as they would need to be to justify the equity premium, low-risk bond yields are far too low. As a result, researchers (e.g. Bansal and Yaron, 2004; Campbell and Cochrane, 1999) have sought to modify the standard CRRA set-up in order to account for these puzzles. In terms of explaining the equity premium and risk-free rate puzzles simultaneously, these models do reasonably well. But they are yet to be examined in terms of their ability to capture the predictability of stock prices with regards to their own history. This predictability, however, is of great practical interest. Given the immense profitability of ex-post predictability, even a small amount of ex-ante predictability could be highly profitable (Farmer and Lo, 1999).

The main type of own-history predictability (i.e. non-zero autocorrelation coefficients) documented is long-horizon mean reversion in stock prices. Mean reversion is the phenomenon by which abnormally high returns are typically followed by abnormally low returns, such that stock prices fluctuate around a trend. Fama and French (1988) and Poterba and Summers (1988) find this predictability to be substantial. While the evidence of mean reversion in stock returns has been questioned, we believe there is still a phenomenon to be explained. Poterba and Summers (1988) and Kim et al. (1991) find that evidence of mean reversion can be sensitive to the time period chosen, in particular to the inclusion of the 1930s and 1940s in the data set. As Section 5 shows, our point estimates do not suggest this. The usual asymptotic standard errors for MDS tests have also been criticised, with several authors (e.g., Kim and Nelson, 1998; Kim et al., 1998; McQueen, 1992; Richardson and Stock, 1990; Richardson and Smith, 1991) arguing they

---

<sup>1</sup>Constant relative risk aversion.

are too small. Nonetheless, Poterba and Summers (1988) show that standard tests for MDS have extraordinarily low power against a mean-reverting alternative. Since even insignificant point estimates often imply reasonably substantial mean reversion, as in our dataset, it is difficult to determine the extent to which a lack of significant mean reversion is merely an expression of this low power.

There has been little recent work on explaining stock return autocorrelations in the context of a rational, consumption-based model. The most recent models do not appear to have been tested in this regard. In order to explain a variety of features of stock returns, including mean reversion, Kandel and Stambaugh (1989) propose adapting the Mehra and Prescott set-up by specifying consumption growth as lognormally distributed with time-varying mean and variance. The mean and variance of consumption growth follow a nine-state Markov-switching process and exhibit positive serial correlation. Kandel and Stambaugh's calibration exercise uses relatively high relative risk aversion (28.6). This shows that the model produces the "U" shaped autocorrelation function observed in stock returns. However, the model is not able to replicate the observed pattern of small positive autocorrelations at short horizons followed by larger negative autocorrelations at longer horizons. Kandel and Stambaugh speculate that this is because their model is overly restrictive. In particular, current news only affects the conditional distribution of consumption one period in the future. Nonetheless, their model broadly matches the observed pattern of autocorrelations at horizons greater than 12 months.

Cecchetti et al. (1990) use a similar specification to Kandel and Stambaugh. Cecchetti et al. use a Markov-switching log endowment level and a more parsimonious two-state specification. They estimate the Markov switching process by maximum likelihood and calibrate the model to US data, using a relative risk-aversion coefficient of 1.7. They find that popular measures of mean reversion always lie within a 60% confidence interval of data simulated from the model. The Cecchetti et al. model has the same problem of not being able to generate negative autocorrelations at short horizons as the Kandel and Stambaugh model. This could be a result of the models being very similar and/or the positive autocorrelations at short horizons being driven by noise. We update the Cecchetti et al. (1990) evidence in two ways. First, we formally estimate their model. This also allows for the development of asymptotic theory for the hypothesis tests used. Second, the Cecchetti et al. (1990) model rests on CRRA preferences. As discussed above, these have been much criticised on an empirical basis, in particular because of the equity premium and risk-free rate puzzles. We test more recent models that can accommodate these two puzzles. Our work also provides an additional specification test of these more recent models, beyond the standard calibration exercises and estimation-based tests which are based on average pricing errors.

Other attempts have been made to explain mean reversion in a risk-based framework. Kim et al. (2001) proxy risk by volatility and use a volatility feedback model.<sup>2</sup> Using a Jegadeesh (1991) test for mean reversion, they find that, adjusting returns for a time-varying expected return where volatility (and therefore expected returns) follows a two-state Markov process, residual returns display no significant mean reversion, despite observed returns doing so. Thus, this simple risk adjustment of returns is sufficient to account for the mean reversion observed. While Ho and Sears (2006) find that Carhart's (1997) extension of the Fama and French (1992) model cannot explain mean reversion, they do not allow the factor loadings to be time-varying. In-keeping with this, Gangopadhyay and Reinganum (1996) find that the CAPM can explain mean reversion only

---

<sup>2</sup>Where an unexpected change in volatility has an immediate impact on stock prices.

when the  $\beta$  (factor loading) can change over time. These adjustments for risk are somewhat crude, and the risk factors are not micro-founded. By contrast, we are testing consumption-based models, which micro-found their risk factors from the start.

There is also a burgeoning literature on how conditional predictability of the short-run equity premium varies with economic and risk conditions.<sup>3,4</sup> Barroso et al. (2017) model the equity risk premium as a function of economic state variables.<sup>5</sup> The extent to which these state variables forecast both the equity risk premium and consumption growth varies with time. When a state variable predicts consumption growth more strongly, it also contributes more to the equity premium. This is entirely consistent with the ICAPM. In other words, a consumption-based asset pricing model appears capable of explaining short-term conditional predictability. Moreover, Baltas and Karyampas (in press) find that the extent to which state variables, such as the dividend yield, can provide out-of-sample predictability of the equity risk premium varies by economic regime.<sup>6</sup> Again, it appears that risk conditions affect the degree of predictability. We extend this work by focussing on specific models of the economic conditions state variables, which are micro-founded, rather than imposed. We consider models with (micro-founded) latent state variables which are conditionally correlated with consumption, and possibly other economic variables, in a time-varying manner. If these latent state variables are what determine risk conditions, this could explain why predictability of the equity premium with respect to observed economic variables appears to vary through time. We also focus on using past returns as the predictor variable and using longer horizons, where returns are generally thought to be more predictable (Campbell, 2014; RSAS, 2013).

This paper proceeds as follows. Section 2 briefly describes and reviews the three asset pricing models considered, Section 3 discusses how we estimate the models and Section 4 discusses the tests we use and how we modify them to account for parameter estimation. Section 5 remarks on the features of our data and reports the estimation of the asset pricing models. Section 6 presents our empirical results regarding the predictability of the model residuals and Section 7 presents our test of a timing strategy to exploit market predictability. Section 8 concludes.

## 2 The models

We consider the ability of the Bansal and Yaron (2004) long-run risk model and the Campbell and Cochrane (1999) habits model to explain mean reversion. We choose these two models as they appear to be the two most influential consumption-based asset pricing models in the literature (see, e.g., the review articles of Cochrane, 2008; Constantinides, 2006; Donaldson and Mehra, 2008; RSAS, 2013; Siegel, 2005). We also consider the Cecchetti et al. (1990) model of equilibrium mean reversion, since this is the only consumption-based model we are aware of specifically developed to account for return predictability.

Below, we review the structure of the models we consider and the evidence regarding

---

<sup>3</sup>Since the risk-free rate is substantially smaller and less variable than the market return (see Section 5), the dynamic properties of the equity risk premium are driven by the market return.

<sup>4</sup>There are also non-risk based explanations for return predictability. These are beyond the scope of this paper.

<sup>5</sup>These are: the dividend yield, default spread, term spread, risk-free rate, price/earnings ratio, value spread, a bond market factor and a term structure factor.

<sup>6</sup>The regimes are: up/down market, upswing/downswing of the business cycle and high/low volatility.

their empirical performance thus far. We see that the Campbell-Cochrane and Bansal-Yaron models are able to justify the market return and risk-free rate in calibrations. The Cecchetti-Lam-Mark model accurately captures stock return autocorrelations in Cecchetti et al.'s (1990) original study.

## 2.1 Bansal-Yaron model

The Bansal and Yaron (2004) model is as follows:

$$V_t = \left[ (1 - \delta)C_t^{1-\frac{1}{\psi}} + \delta (E_t [V_{t+1}^{1-\gamma}])^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \quad (1)$$

$$x_{t+1} = \rho_x x_t + \psi_x \sigma_t \varepsilon_{t+1} \quad (2)$$

$$\Delta c_{t+1} = \mu_c + x_t + \sigma_t \eta_{t+1} \quad (3)$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} \quad (4)$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1} \quad (5)$$

$$\varepsilon_t, \eta_t, u_t, w_t \sim NID(0, 1),$$

where  $V_t$  is the representative investor's value function,  $\delta$  the subjective discount factor,  $\gamma > 0$  the risk-aversion coefficient,  $\psi > 0$  the elasticity of intertemporal substitution (*EIS*),  $C_t$  consumption,  $D_t$  dividends,  $E_t$  the conditional expectation at time  $t$  and lower-case variables denote logs of upper-case variables.

The model has three key ingredients. The first is recursive preferences (1) à la Epstein and Zin (1989) and Weil (1989). Unlike standard CRRA preferences, these allow *EIS* and risk aversion to differ. This represents a methodological advantage since risk aversion and intertemporal substitution are different concepts. *EIS* reflects the extent to which consumers are willing to smooth certain consumption through time, while risk aversion relates to the extent to which consumers are willing to smooth consumption across uncertain states of nature (Cochrane, 2008).

The second ingredient is to specify consumption growth (3) as having a small predictable component (the long-run risk,  $x$ ) so that consumption news in the present affects expectations of future consumption growth. By making consumption growth persistent, the long-run risk can increase the impact of present consumption growth on the difference between present discounted values of dividend streams and therefore returns. It is the differences between these PDVs that ultimately drive returns in consumption-based models. There is great debate as to whether or not it is reasonable to model consumption growth as not being *iid* (e.g., Cochrane, 2008). While this matter is beyond our scope, there are plausible reasons for consumption growth not being *iid*, such as binding credit constraints and precautionary savings.

The third ingredient is allowing for time-varying economic volatility (5) in consumption growth. This reflects time-varying economic uncertainty and is a further source of investor uncertainty and risk.

Bansal and Yaron calibrate their model on a monthly basis such that it matches features of annualised US data. They show their model can justify the equity premium, risk-free rate and the volatilities of the market return, risk-free rate and price-dividend ratio. Moreover, the price-dividend ratio predicts returns and the volatility of returns is time-varying, as in the data. When *EIS*  $> 1$ , the model also captures leverage effects

(a rise in volatility being associated with a stock market fall) and the observed negative correlation between consumption volatility and price-dividend ratios. Time-varying economic uncertainty adds a further risk factor and allows the equity premium to vary through time. The reliance of some of the calibration results on  $EIS > 1$  may cause some concern.  $EIS$  estimates are not conclusively greater or less than one. Bansal and Yaron present some empirical evidence suggesting  $EIS > 1$ . They argue that forcing consumption volatility to be constant, as most estimates of  $EIS$  that are less than one do, biases the  $EIS$  estimated downwards.

Constantinides and Ghosh (2011) estimate the Bansal-Yaron model by GMM. The results are mixed. The model matches the unconditional moments of aggregate dividend and consumption growth rates. As predicted, the long-run risk forecasts consumption and dividend growth. However, the implied risk-free rate is too high and insufficiently variable. The market price-dividend ratio is not sufficiently variable and the model requires greater persistence in consumption and dividend growth than is observed. At long horizons, the conditional variance does not forecast the equity premium but does forecast consumption and dividend growth, contrary to the model's predictions. Moreover, the  $J$ -statistic  $p$ -value is less than 0.03 in all specifications considered.

## 2.2 Campbell-Cochrane model

Campbell and Cochrane's (1999) model adds a slow-moving external habit to the standard power utility function. The representative agent's utility function is

$$U_t(C) = E_t \sum_{s=0}^{\infty} \delta^s \frac{(C_{t+s} - H_{t+s})^{1-\gamma} - 1}{1-\gamma},$$

where  $\delta$  is the subjective discount factor,  $\gamma$  the utility curvature and  $H_t$  the habit level of consumption. Defining  $S_t \equiv (C_t - H_t)/C_t$ , the habit evolves according to

$$s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)\nu_{t+1}, \quad (6)$$

where  $\bar{s}$  is the steady-state  $s$  and  $\bar{S} = \sigma_\nu \sqrt{\gamma/(1-\phi)}$ . Consumption and dividends are such that

$$\begin{aligned} \Delta c_t &= \bar{g} + \nu_t \\ \Delta d_t &= \bar{g} + w_t \end{aligned}$$

with  $\Delta$  being the first difference operator and

$$\begin{pmatrix} \nu_t \\ w_t \end{pmatrix} \sim NID \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\nu^2 & \sigma_{\nu w} \\ \sigma_{\nu w} & \sigma_w^2 \end{pmatrix} \right).$$

$\lambda(s_t)$  is a sensitivity function given by

$$\lambda(s_t) = \begin{cases} (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1, & \text{if } s_t \leq s_{\max} \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

with  $s_{\max} \equiv \bar{s} + \frac{1}{2}(1 - \bar{S}^2)$ . Campbell and Cochrane set  $\phi$  to be equal to the first-order autocorrelation coefficient of the log market price-dividend ratio,  $z_{m,t}$ .

Campbell and Cochrane calibrate their model to match the annualised unconditional equity premium using monthly US data. When given actual data, the model replicates the main movements observed in stock prices. In simulations, the model is able to justify the means and standard deviations of excess returns and the price-dividend ratio, and the existence of a short-run and long-run equity premium. Moreover, this is achieved without a risk-free rate puzzle by construction: the habit is specified such that the risk-free rate remains constant and the model is calibrated such that the log risk-free rate is equal to the sample mean.<sup>7</sup> In Garcia et al.’s (2004) GMM estimation of the Campbell-Cochrane model, the estimated  $\gamma$  is significantly greater than 0 and the  $\delta$  significantly less than 1. The  $J$ -statistic  $p$ -value exceeds 0.2, although this does condition on earlier estimates of time-series parameters.

One may worry about the habit being external, not internal. However, Campbell and Cochrane (1998) show that the model’s predicted asset prices are unaffected when the habit is internal, habit accumulation linear and the risk-free rate constant. The only change is that log consumption growth must now be specified as only almost, but not exactly, *iid*.

Another concern is that the model implies a high average relative risk aversion (Salomons, 2008). However, as Campbell and Cochrane (1999) note, high implied risk aversion is a problem in terms of the model matching microeconomic evidence. For medium and large-stakes gambles of wealth, evidence is mainly based on surveys or the author’s intuition. Campbell and Cochrane argue that, since these gambles are larger than what consumers are used to dealing with, they may struggle to comprehend the situation in question and so this evidence is not necessarily conclusive. As already shown, the model works well from a macroeconomic perspective.

A final concern is that utility is undefined for consumption less than habit. However, the habit is specified such that  $C_t > H_t$  always. Yet, this specification means the average habit is only around 5% less than average consumption, making habits seem extreme. This is a minor concern that Campbell and Cochrane suggest could be resolved by accounting for concerns about extreme losses, sampling variation and survivorship bias in their calibration.

## 2.3 Cecchetti-Lam-Mark model

Cecchetti et al.’s (1990) model attempts to explain mean reversion in a rational framework. The model is an endowment economy where the representative consumer has CRRA preferences:

$$U_t(C) = E_t \sum_{s=0}^{\infty} \delta^s \frac{C_{t+s}^{1-\gamma} - 1}{1-\gamma}.$$

Here,  $\delta$  denotes the subjective discount factor and  $\gamma$  the coefficient of relative risk aversion. Taking (log) consumption as the appropriate endowment process,

$$c_{t+1} = c_t + \alpha_0 + \alpha_1 y_t + \varepsilon_{t+1}. \tag{8}$$

$y_t \in \{0, 1\}$  is a first-order Markov process and  $\varepsilon_t \sim NID(0, \sigma^2)$ .  $y_t = 1$  denotes a “bust” period, so  $\alpha_1$  is restricted to be less than zero.

---

<sup>7</sup>Campbell and Cochrane argue this is realistic as the risk-free rate varies relatively little and does not vary cyclically.

Cecchetti et al. (1990) test their specification against annual US data in terms of how well their estimated Markov process matches the endowment process and how well the mean reversion of their equilibrium asset prices matches that observed in reality. Mean reversion is examined in terms of variance ratios and Fama and French (1988) regression coefficients using annual US/S&P data over 2-10 year horizons. Tests of how well the model explains mean reversion are carried out by simulating 10,000 “small” samples (of length 116, the length of their data) and 10,000 large samples (length 1160) of returns, both for linear and concave utility.<sup>8</sup> They then compute the variance ratios and regression coefficients from these sample, thus forming Monte Carlo distributions of the statistics. Their observed values lie inside the 60% confidence intervals implied by both the linear and concave utility models. However, the small-sample medians of the statistics from the concave utility model are somewhat closer to those observed than those from the linear utility model. Cecchetti et al. conclude that the concave utility model better explains mean reversion. The Markov bust-state probability is realistic at and, based on 1,000 replications of the estimated Markov process, the sample moments of the data are fairly close to those of the Markovian endowment.

The main concern regarding the Cecchetti-Lam-Mark model is that, given that it uses standard CRRA preferences, it seems highly likely to suffer from the risk-free rate and equity premium puzzles.<sup>9</sup> Indeed, Cecchetti et al. note that their model was designed to explain only mean reversion in asset prices, not asset prices in general. However, given the empirical success of this model in explaining mean-reversion, it remains a useful benchmark for our analysis.

There is one concern that affects all three models. They are all representative-agent models and neglect investor heterogeneity.<sup>10</sup> The subset of the population owning stocks may face a different consumption process to the general population. Correcting for this requires access to panel data for consumption and investment holdings. These data do not go back far enough to be able to estimate medium and long-horizon return statistics with a reasonable sample size. This lack of data rules out testing models such as the Constantinides et al. (2002) “junior can’t borrow” model and Malloy et al.’s (2009) model of heterogeneous investors. Given the data available, this concern must be put to one side and left for future research.

### 3 Model estimation

We estimate the asset pricing models by GMM, as detailed below, using quarterly data. We use quarterly data to increase the sample size, while maintaining a sufficiently long horizon such that one would expect consumption and investment decisions to be correlated.

#### 3.1 Bansal-Yaron model

Constantinides and Ghosh (2011) show that the log-linearised version of the Bansal-Yaron model can be inverted, such that the unobserved state variables can be written as a linear

---

<sup>8</sup>Concave utility is assumed to exhibit relative risk aversion of 1.7, which is rather low.

<sup>9</sup>Although, strictly speaking, there is no risk-free rate in the Cecchetti-Lam-Mark model.

<sup>10</sup>Campbell and Cochrane (1999) show that their model can admit some forms of investor heterogeneity. These are too limited to be of much use.



combination of observables. In particular,

$$x_t = \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 z_{m,t} \quad (9)$$

$$\sigma_t^2 = \beta_0 + \beta_1 r_{f,t} + \beta_2 z_{m,t} \quad (10)$$

where  $\alpha_0, \dots, \beta_2$  are functions of Bansal-Yaron model parameters, as detailed in Appendix A.1, and  $r_{f,t}$  the (log) risk-free rate. This allows them to express the Bansal-Yaron Euler equation for a general asset as

$$\mathbb{E}_t \left[ \exp \left\{ a_1 + a_2 \Delta c_{t+1} + a_3 \left( r_{f,t+1} - \frac{1}{\kappa_1} r_{f,t} \right) + a_4 \left( z_{m,t+1} - \frac{1}{\kappa_1} z_{m,t} \right) + r_{t+1} \right\} \right] - 1 = 0,$$

where  $r_t$  is the log asset return and  $a_1, \dots, a_4, \kappa_1$  are functions of the Bansal-Yaron model parameters, also given in Appendix A.1.

In addition, they derive nine unconditional moment restrictions for continuously compounded consumption and dividend growth, which are given in Appendix A.2. These moment conditions are derived from Bansal and Yaron's (2004) specification of consumption and dividend growth, the long-run risk and its conditional variance.

The model has 12 parameters to estimate. We use an Euler equation for each asset considered, and the nine time-series restrictions as our GMM moment conditions.

Constantinides and Ghosh show that

$$\mathbb{E}_t r_{m,t+1} = B_0 + B_1 x_t + B_2 \sigma_t^2$$

where  $r_{m,t}$  is the market return and  $B_0, \dots, B_2$  are non-linear combinations of the 12 model parameters provided in Appendix A.3. It follows that

$$\mathbb{E}_t r_{m,t+q+1} = B_0 + B_1 \rho_x^q x_t + B_2 (\sigma^2 (1 - (1 - \nu)^q) + \nu^q \sigma_t^2).$$

By (9)-(10), we have expected future returns in terms of observables up to  $t$ . These yield a plug-in estimator of  $\mathbb{E}_t r_{m,t+q+1}$ .

### 3.2 Campbell-Cochrane model

We use a GMM procedure along the lines of that of Garcia et al. (2004) for estimating the Campbell-Cochrane model. The procedure has three parts. First, the time-series parameters  $\bar{g}$ ,  $\sigma_\nu^2$  and  $\sigma_w^2$  are estimated by GMM, using the following four moment conditions:

$$\begin{aligned} 0 &= \mathbb{E}(\Delta c_{t+1} - \bar{g}) \\ 0 &= \mathbb{E}(\Delta c_{t+1} - \bar{g})^2 - \sigma_\nu^2 \\ 0 &= \mathbb{E}(\Delta d_{t+1} - \bar{g}) \\ 0 &= \mathbb{E}(\Delta d_{t+1} - \bar{g})^2 - \sigma_w^2. \end{aligned}$$

Second, we estimate  $\alpha$  and  $\phi$  from the linear regression

$$z_{m,t+1} = \alpha + \phi z_{m,t} + e_{t+1}.$$

Third, we estimate the preference parameters  $\delta$  and  $\gamma$ . Since the Euler equation in the Campbell-Cochrane model is given by

$$\mathbb{E}_t \left[ \delta \left( \frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma} (1 + R_t) \right] - 1 = 0, \quad (11)$$

it is first necessary to generate the series  $s_t$ . We do this by initialising the series at  $s_0 = \bar{s} = \ln(\sigma_\nu \sqrt{\gamma/(1-\phi)})$ , using the estimates of the relevant time-series moments from above and assuming an initial  $\gamma$  of 2. This allows the series  $s_t$  to be generated as per (6) and (7). We then use an Euler equation of the form (11) for each asset (excluding the risk-free rate) to make up the moment conditions for estimating  $\delta$  and  $\gamma$ . This estimate of  $\gamma$  is then used in place of the assumed value of two to generate a new  $s_t$  series and the preference parameters are estimated again by GMM. We iterate this procedure until the estimates of  $\delta$  and  $\gamma$  converge.<sup>11</sup>

We estimate  $E_t r_{m,t+q+1}$  the Campbell-Cochrane model as follows. By rearranging and iterating the Euler equation forwards, we have

$$P_t = \sum_{j=1}^{\infty} \delta^j E_t \left[ \left( \frac{S_{t+j} C_{t+j}}{S_t C_t} \right)^{-\gamma} D_{t+j} \right] \quad (12)$$

when we impose the no-bubble condition

$$\lim_{j \rightarrow \infty} \delta^j E_t \left[ \left( \frac{S_{t+j} C_{t+j}}{S_t C_t} \right)^{-\gamma} P_{t+j} \right] = 0.$$

Therefore,

$$E_t(1 + R_{t+q}) = E_t \frac{\sum_{j=1}^{\infty} \delta^j E_{t+q} \left[ \left( \frac{S_{t+q+j} C_{t+q+j}}{S_{t+q} C_{t+q}} \right)^{-\gamma} D_{t+q+j} \right]}{\sum_{j=1}^{\infty} \delta^j E_{t+q-1} \left[ \left( \frac{S_{t+q+j-1} C_{t+q+j-1}}{S_{t+q-1} C_{t+q-1}} \right)^{-\gamma} D_{t+q+j-1} \right]}. \quad (13)$$

Since  $R_t \approx r_t$ , we have an estimator of  $E_t r_{m,t+q+1}$  when it is the market dividends that are simulated.

We estimate (13) by means of a “double” simulation. First, we simulate the series  $c_{t+1}, \dots, c_{t+q}$ ,  $d_{t+1}, \dots, d_{t+q}$  and  $s_{t+1}, \dots, s_{t+q}$ , conditioning on  $c_t$ ,  $d_t$  and  $s_t$ . Second we simulate  $c_{t+q}, \dots, c_{t+q+99}$ ,  $d_{t+q}, \dots, d_{t+q+99}$  and  $s_{t+q}, \dots, s_{t+q+99}$ , and  $c_{t+q-1}, \dots, c_{t+q+98}$ ,  $d_{t+q-1}, \dots, d_{t+q+98}$  and  $s_{t+q-1}, \dots, s_{t+q+98}$  separately for the numerator and the denominator respectively.<sup>12</sup> The expectation in the numerator in (13) is then estimated by the mean of 200 simulations of the PDV and likewise for the denominator, such that the ratio of the expectations can be computed. The “first-stage” simulation (between  $t$  and  $t+q$ ) is repeated 200 times and the expectation of the ratio is estimated by the mean of the simulated ratios. Our simulation procedure is shown schematically in Figure 1.

### 3.3 Cecchetti-Lam-Mark model

We use to estimate  $\delta$  and  $\gamma$ . The moment conditions comprise an Euler equation for each asset of the form

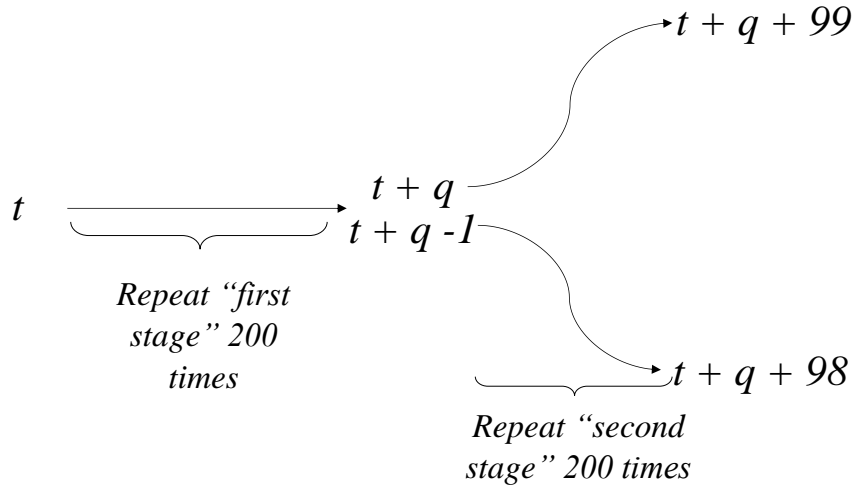
$$E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + R_t) \right] - 1 = 0. \quad (14)$$

We estimate the Markov switching endowment process by maximum likelihood (Hamilton, 1989). Hamilton also provides algorithms for extracting smoothed and filtered probabilities of each state. In a slight deviation from Cecchetti et al. (1990), we estimate a Markov-switching process where the consumption innovation  $\varepsilon_{t+1} | y_t \sim N(0, \sigma_{y_t}^2)$ .

<sup>11</sup>Our tolerance is a difference of up to  $10^{-7}$  between the estimates of different iterations.

<sup>12</sup>We truncate the horizon simulated to be 99 quarters into the future.

Figure 1: Schematic representation of “two-step” simulation



The arrows represent simulation. We first simulate from  $t$  to  $t+q$  and then separately from  $t+q$  to  $t+q+99$  for the numerator and from  $t+q-1$  to  $t+q+98$  for the denominator. The mean of the 200 second-stage ratios gives one observation of the first-stage return. The mean of the 200 first-stage return gives the expected return.

Cecchetti et al. show that

$$1 + R_{m,t+1} = \frac{1 + \kappa(y_{t+1})}{\kappa(y_t)} \exp\{\alpha_0 + \alpha_1 y_t + \varepsilon_t\} \quad (15)$$

where  $\kappa(y_t)$  is a non-linear function of model parameters defined in Appendix B. We evaluate  $E_t R_{m,t+q+1}$  for  $q > 0$  by simulating a series  $y_{t+1}, \dots, y_{t+q+1}$  given  $y_t$ . The expectation is then taken over 200 simulated returns.

## 4 Tests

To test whether the asset-pricing models discussed above capture the serial correlation structure of stock returns, we note that rational expectations imply

$$r_{t+1} = E_t r_{t+1} + \xi_{t+1}, \quad (16)$$

where expectations are formed under the model in question and  $\xi_{t+1}$  is unforecastable at  $t$ . If the model accurately captures own-history predictability,  $\xi_t$  should be MDS. If not, there is clearly something in the own-history predictability structure of  $r_t$  not captured by  $E_{t-1} r_t$ . We estimate (16) using the plug-in estimators of  $E_t r_{t+1}$  derived above and base our tests on the resulting residual  $\hat{\xi}_t$ , where  $\hat{\xi}_t$  is re-centred such that it has mean zero. As predictability is typically considered a longer-term phenomenon, the return horizons considered in each test are  $q \in \{2, 3, \dots, 10\}$  years, i.e.  $q \in \{8, 24, \dots, 40\}$  quarters. We consider tests of linear and non-linear predictability, as well as a rescaled range test. In each case, we adapt the test to cope with the fact that  $\mu_t(\boldsymbol{\theta}) \equiv E_{t-1} r_{m,t}$  is estimated and this estimate,  $\mu_t(\hat{\boldsymbol{\theta}})$ , is a function of a parameter vector estimated by GMM. Poterba

and Summers (1988) demonstrate the low power of tests of the martingale difference hypothesis based on serial correlation coefficients against a persistent but negatively serially correlated alternative. Given their arguments, we use a 10% significance level throughout, although we also report  $p$ -values.

## 4.1 Linear predictability

We use Lo and MacKinlay's (1988) variance-ratio test, which exploits the fact that, for serially uncorrelated  $\xi_t$ ,  $\text{Var}(\sum_{s=1}^q \xi_{t+s}) = q \text{Var}(\xi_t)$ . Poterba and Summers (1988) and Lo and MacKinlay (1989) show this is generally a more powerful test of the martingale difference hypothesis than unit root and autoregressive tests. The population variance ratio is given by  $VR(q) = \text{Var}(\sum_{s=1}^q \xi_{t+s})/q \text{Var}(\xi_t)$ . Cochrane (1988) shows that

$$VR(q) = 1 + 2 \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \rho(j), \quad (17)$$

where  $\rho(j)$  is the  $j^{\text{th}}$  order serial correlation coefficient of  $\xi_t$ . Thus  $VR(q) < 1$  can be taken as evidence of negative serial correlation in returns dominates at horizon  $q$ , and vice-versa for  $VR(q) > 1$ . Under the null that  $\xi_t$  is MDS,  $VR(q) = 1$ .

We estimate (17) by replacing  $\rho(j)$  with  $\tilde{\rho}(j)$ , where

$$\begin{aligned} \tilde{\rho}(j) &= \frac{1}{T} \sum_{t=j+1}^T \left( r_{m,t} - \mu_t^*(\hat{\boldsymbol{\theta}}) \right) \left( r_{m,t-j} - \mu_{t-j}^*(\hat{\boldsymbol{\theta}}) \right) \\ \mu_t^*(\boldsymbol{\theta}) &= \mathbb{E}_{t-j-1}(\mu_t(\boldsymbol{\theta})). \end{aligned}$$

This approach corrects for the estimation of  $\hat{\boldsymbol{\theta}}$ . To see this, consider the asymptotic second-order expansion of  $\tilde{\rho}(j)$ :

$$\begin{aligned} \tilde{\rho}(j) &= \frac{1}{T} \sum_{t=j+1}^T (r_{m,t} - \mu_t^*(\boldsymbol{\theta}))(r_{m,t-j} - \mu_{t-j}^*(\boldsymbol{\theta})) - (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \mathbb{E} \left[ \frac{\partial \mu_t^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (r_{m,t-j} - \mu_{t-j}^*(\boldsymbol{\theta})) \right] \\ &\quad - (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \mathbb{E} \left[ (r_{m,t} - \mu_t^*(\boldsymbol{\theta})) \frac{\partial \mu_{t-j}^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] + o_P(T^{-1}). \end{aligned} \quad (18)$$

Under the null that  $\xi_t$  is an MDS,  $\mathbb{E} \left[ (r_{m,t} - \mu_t^*(\boldsymbol{\theta})) \frac{\partial \mu_{t-j}^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] = 0$  and  $\mathbb{E} \left[ \frac{\partial \mu_t^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (r_{m,t-j} - \mu_{t-j}^*(\boldsymbol{\theta})) \right] = 0$ . Therefore, asymptotically,

$$\tilde{\rho}(j) = \frac{1}{T} \sum_{t=j+1}^T (r_{m,t} - \mu_t^*(\boldsymbol{\theta}))(r_{m,t-j} - \mu_{t-j}^*(\boldsymbol{\theta})).$$

Thus Lo and MacKinlay's (1988) asymptotics can be applied. In particular, our estimated  $VR(q)$  is asymptotically normal. Lo and MacKinlay show that  $\tau^*(q) \stackrel{a}{\sim} N(0, 1)$ , where

$\tau^*(q) \equiv \sqrt{T}(\widehat{VR}(q) - 1)/\sqrt{\widehat{\lambda}(q)}$  and

$$\begin{aligned}\widehat{\lambda}(q) &\equiv \sum_{j=1}^{q-1} \left[ \frac{2(q-j)}{q} \right]^2 \widehat{\vartheta}(j) \\ \widehat{\vartheta}(j) &\equiv \frac{\sum_{k=j+1}^T \left( r_k - \mu_k^*(\widehat{\boldsymbol{\theta}}) \right)^2 \left( r_{k-j} - \mu_{k-j}^*(\widehat{\boldsymbol{\theta}}) \right)^2}{\left[ \sum_{k=j+1}^T \left( r_k - \mu_k^*(\widehat{\boldsymbol{\theta}}) \right)^2 \right]^2}.\end{aligned}$$

These results allow for heteroscedasticity and  $\tau^*(q)$  is the statistic we base our inference on.<sup>13</sup>

## 4.2 Non-linear predictability

The variance-ratio statistic is a function of the sample autocorrelations of  $\widehat{\xi}_t$  and therefore does not exploit the full hypothesised MDS structure of  $\xi_t$ . In particular it neglects non-linear predictability. We test for non-linear predictability using Linton and Whang's (2007) quantilogram, which is based on the correlation of quantile hits. If  $\widehat{\xi}_t$  is MDS, the probability  $\widehat{\xi}_{t+k}$  is in the  $\alpha$  quantile given  $\widehat{\xi}_t$  is in the  $\alpha$  quantile should remain  $\alpha$ . The quantile hits are uncorrelated. The quantilogram is a more general version of Wright's (2000) sign tests, which focus on whichever quantile zero is in.

In our test statistic, we weight the quantilogram estimates, analogously to the variance ratios. This gives

$$\widehat{W}_\alpha(q) = \sum_{j=1}^{q-1} \left( 1 - \frac{j}{q} \right) \widehat{\rho}_\alpha(j), \quad (19)$$

where

$$\begin{aligned}\widehat{\rho}_\alpha(j) &= \frac{\sum_{t=1}^{T-j} \psi_\alpha(\widehat{\xi}_t - \widehat{\mu}_\alpha) \psi_\alpha(\widehat{\xi}_{t+j} - \widehat{\mu}_\alpha)}{\sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\widehat{\xi}_t - \widehat{\mu}_\alpha)} \sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\widehat{\xi}_{t+j} - \widehat{\mu}_\alpha)}} \\ \psi_\alpha(\cdot) &= \alpha - \mathbb{1}(\cdot < \alpha) \\ \widehat{\mu}_\alpha &= \operatorname{argmin}_{m \in \mathbb{R}} \sum_{t=1}^T (\widehat{\xi}_t - m) \times \psi(\widehat{\xi}_t - m).\end{aligned}$$

and  $\mathbb{1}(\cdot)$  is the indicator function. We evaluate (19) over the same  $q$  as in the variance ratio tests and over a range of both extreme and moderate quantiles, namely  $\alpha \in \{0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$ .

Inference is carried out using a bootstrap. We apply a wild bootstrap to the estimated residuals,  $\widehat{\xi}_t$ . That is to say,  $\widehat{\xi}_t$  is pre-multiplied by  $\iota_t^*$  at each  $t$ , where  $E(\iota_t^*) = 0$  and

<sup>13</sup>To see why it is necessary to use  $\mu_t^*(\boldsymbol{\theta})$  in place of  $\mu_t(\boldsymbol{\theta})$  consider the asymptotic expansion of  $\widehat{\rho}(j) = \frac{1}{T} \sum_{t=j+1}^T \left( r_{m,t} - \mu_t(\widehat{\boldsymbol{\theta}}) \right) \left( r_{m,t-j} - \mu_{t-j}(\widehat{\boldsymbol{\theta}}) \right)$ , which is analogous to (18). While, under the null,  $E \left[ \left( r_{m,t} - \mu_t(\boldsymbol{\theta}) \right) \frac{\partial \mu_{t-j}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] = 0$ , this is not generally true for  $E \left[ \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left( r_{m,t-j} - \mu_{t-j}(\boldsymbol{\theta}) \right) \right]$ . Lo and MacKinlay's (1988) asymptotic results, which are based on only the first term in (18), do not necessarily go through.

$\text{Var}(\iota_t^*) = 1$ . In this case,  $\iota_t^*$  follows Mammen's (1993) two-point distribution.<sup>14</sup> Then, we use the bootstrapped residuals to extract a pseudo-sample of returns  $r_{m,t}^*$  by the relationship

$$r_{m,t}^* = E_{t-1} r_{m,t} + \iota_t^* \hat{\xi}_t.$$

We use  $r_{m,t}^*$  to generate a new series for the market value and therefore obtain the pseudo-sample of the log price-dividend ratio,  $z_{m,t}^*$ . This modified data set is then used to re-estimate the asset pricing model parameters and, in turn, generate a pseudo-sample of expected returns and thus a (new) pseudo-sample of residuals. It is this final pseudo-sample of residuals which we use to calculate our weighted quantilogram. The empirical distribution of the weighted quantilograms thus obtained is used for inference and the bootstrap procedure is repeated 200 times. Notice that our procedure accounts for the estimation of  $E_{t-1} r_{m,t}$  and conditions on consumption and dividends.

### 4.3 Rescaled range

In addition to the above tests of the null that  $\xi_t$  is MDS, we consider a rescaled range test. We do so as the rescaled range may be a more powerful test of the martingale difference hypothesis than those above in the presence of long-range dependence (Lo, 1991). Our test statistic is novel as it accounts for the estimation of  $\theta$ . Our test statistic  $\hat{Q}^K$  is given by

$$\hat{Q}^K = \frac{1}{\hat{s}\sqrt{T}} \left[ \max_{1 \leq k \leq T} \sum_{t=1}^k \hat{K}_t - \min_{1 \leq k \leq T} \sum_{t=1}^k \hat{K}_t \right] \quad (20)$$

$$\hat{K}_t = r_t - \mu_t(\hat{\theta}) + \dot{\mu}_t(\hat{\theta}) \left[ \sum_{k=1}^{t-1} \dot{\mu}_k(\hat{\theta})(\dot{\mu}_k(\hat{\theta}))' \right]^{-1} \sum_{k=1}^{t-1} \dot{\mu}_k(\hat{\theta}) (r_{m,k} - \mu_k(\hat{\theta})) \quad (21)$$

$$\hat{s}^2 = \frac{1}{T} \sum_{t=1}^T \hat{\xi}_t^2. \quad (22)$$

To adapt the rescaled range test to cope with the estimation of  $\theta$ , we start with Hurst's (1951) classic rescaled range statistic,  $\hat{Q}$ , which has

$$\hat{Q} = \frac{1}{\hat{s}\sqrt{T}} \left[ \max_{1 \leq k \leq T} \sum_{t=1}^k (r_{m,t} - \mu_t(\hat{\theta})) - \min_{1 \leq k \leq T} \sum_{t=1}^k (r_{m,t} - \mu_t(\hat{\theta})) \right].$$

Noting that  $\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \Omega)$  for some  $\Omega$ , we write

$$\begin{aligned} \mu_t(\hat{\theta}) &= \mu_t(\theta) + \dot{\mu}_t(\theta)(\hat{\theta} - \theta) + o_P(T^{-1/2}) \\ \dot{\mu}_t(\theta) &= \frac{\partial \mu_t(\theta)}{\partial \theta}. \end{aligned}$$

<sup>14</sup> $\iota_t^*$  is *iid* through time and has probability mass function

$$f_I(\iota_t^*) = \begin{cases} \frac{\sqrt{5}+5}{10}, & \iota_t^* = \frac{1-\sqrt{5}}{2} \\ \frac{\sqrt{5}-5}{10}, & \iota_t^* = \frac{1+\sqrt{5}}{2} \end{cases}$$

We can therefore first approximate  $\widehat{Q}$  by

$$\begin{aligned} Q &= \frac{1}{\sigma\sqrt{T}} \left[ \max_{1 \leq k \leq T} \sum_{t=1}^k \left( r_t - \mu_t(\boldsymbol{\theta}) - \dot{\mu}_t(\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \right) - \min_{1 \leq k \leq T} \sum_{t=1}^k \left( r_t - \mu_t(\boldsymbol{\theta}) - \dot{\mu}_t(\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \right) \right] \\ &= \max_{1 \leq k \leq T} \left( \frac{1}{\sigma\sqrt{T}} S_k - \frac{1}{T\sigma} T_k \times \Omega^{1/2} \times Z \right) - \min_{1 \leq k \leq T} \left( \frac{1}{\sigma\sqrt{T}} S_k - \frac{1}{T\sigma} T_k \times \Omega^{1/2} \times Z \right) \end{aligned}$$

where  $S_k = \sum_{t=1}^k (r_t - \mu_t(\boldsymbol{\theta}))$ ,  $T_k = \sum_{t=1}^k \dot{\mu}_t(\boldsymbol{\theta})$  and  $\Omega^{1/2}Z$  is the limiting random variable of the sequence  $\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$  and  $Z \sim N(0, I)$ . Replacing  $T_k$  by  $k \mathbb{E}(\dot{\mu}_t(\boldsymbol{\theta}))$ ,

$$Q = \sup_{u \in [0,1]} \left( B(u) - \sigma^{-1} u \mathbb{E}(\dot{\mu}_t(\boldsymbol{\theta})) \times \Omega^{1/2} \times Z \right) - \inf_{u \in [0,1]} \left( B(u) - \sigma^{-1} u \mathbb{E}(\dot{\mu}_t(\boldsymbol{\theta})) \times \Omega^{1/2} \times Z \right).$$

The Khamaladze transform (Brown et al., 1975) for the rescaled range process is then given by (20)-(22).

Since we have now corrected the rescaled range process to account for the estimation of  $\boldsymbol{\theta}$ ,  $\widehat{Q}^K$  has the same pivotal asymptotic distribution under the null that  $\xi_t$  is MDS dependence as  $\widehat{Q}$  does. This distribution is obtained by Lo (1991) as

$$F_Q(v) = 1 + 2 \sum_{k=1}^{\infty} (1 - 4k^2 v^2) \exp\{-2(kv)^2\}.$$

#### 4.4 Maximal predictability

Huang and Zhou (2015) develop a Wald test of whether the predictability of excess returns,  $\tilde{r}_{t+1} = r_{t+1} - r_{f,t+1}$ , is too large. Predictability is measured with respect to a forecasting variable,  $f_t$ . “Too large” is defined as too large to be consistent with  $\widetilde{M}_t$ , the stochastic discount factor (SDF) normalised such that  $\mathbb{E} \widetilde{M}_{t+1} = 1$ , being a function of a given set of state variables  $\boldsymbol{\omega}_t$ .<sup>15</sup> The test is semi-parametric in that the functional form of the SDF need not be known. The Wald statistic tests whether theoretical upper bound on  $R^2$  implied by the state variables is exceeded by the empirical  $R^2$  from the univariate one-step-ahead predictive regression of  $\tilde{r}_{t+1}$  on  $f_t$ . It is straightforward to verify that this test applies almost directly to the  $q$ -step-ahead predictive regression

$$\tilde{r}_{t+q} = \alpha + \beta f_t + \varepsilon_{t+q}.$$

In this context, when bounding  $R^2$  with  $\text{SR}(r_m)$ , the market Sharpe ratio,<sup>16</sup> the bound becomes

$$R^2 \leq \bar{R}_{SR}^2 = \phi_{\omega, r_f}^2 h^2 \text{SR}^2(r_m),$$

<sup>15</sup>Our other tests relate to actual, not excess returns. However, as mentioned earlier,  $r_{f,t}$  is substantially smaller and less variable than  $r_{m,t}$ . Thus the dynamic properties of  $\tilde{r}_t$  are driven by  $r_{m,t}$ .

<sup>16</sup>We choose to bound  $R^2$  with the market Sharpe ratio, as opposed to risk aversion. The latter requires specifying a maximum admissible risk aversion, which is undesirable for reasons explained by Campbell and Cochrane (1999). Moreover, the functional form of the risk-aversion bound requires the market portfolio to be the optimal wealth and lognormally distributed. This is not an assumption we wish to make. Nonetheless, the choice of bound makes little difference to Huang and Zhou’s (2015) results due to their specifications of  $h$  and the maximum risk aversion.

where

$$\phi_{\omega,rf}^2 = \rho_{\omega,rf}^2 \frac{\text{Var}[\tilde{r}_{t+q}(\tilde{r}_t - \mu_f)]}{\text{Var}(\tilde{r}_{t+q}) \text{Var}(f_t)}$$

$$\rho_{\omega,rf}^2 = \frac{\text{Cov}[\boldsymbol{\omega}_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]' \text{Var}^{-1}(\boldsymbol{\omega}_{t+q}) \text{Cov}[\boldsymbol{\omega}_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]}{\text{Var}[\tilde{r}_{t+q}(f_t - \mu_f)]},$$

and  $\mu_f = E(f_t)$ .  $h$  is a parameter chosen by the marginal investor. We follow Cochrane and Saá-Requejo (2000) in using  $h = 2$ . This bound requires  $\boldsymbol{\omega}$  to have an elliptic distribution. Huang and Zhou argue assuming an elliptic distribution works well for the state variables of the Bansal-Yaron and Campbell-Cochrane models.

Huang and Zhou's test exploits the asymptotic normality of standard estimators of the mean and covariance matrix of  $(r_{t+q}, f_t, r_{t+q}f_t, \boldsymbol{\omega}'_{t+q})'$ . These means and covariances, which comprise  $\boldsymbol{\theta}_{SR}$ , are all that is required to calculate the empirical  $R^2$  and its bound. We estimate  $\boldsymbol{\theta}_{SR}$  by GMM.

Testing whether  $R^2$  exceeds  $\bar{R}_{SR}^2$  is equivalent to a one-sided test of the null  $f(\boldsymbol{\theta}_{SR}) \equiv R^2 - \bar{R}_{SR}^2 = 0$  against the alternative that  $f(\boldsymbol{\theta}_{SR}) \equiv R^2 - \bar{R}_{SR}^2 > 0$ . The Wald statistic for this test is

$$W_{RA} = T f(\hat{\boldsymbol{\theta}}_{SR}) \left[ \frac{df}{d\boldsymbol{\theta}_{SR}} \text{Var}(\hat{\boldsymbol{\theta}}_{SR}) \frac{df}{d\boldsymbol{\theta}_{SR}} \right]^{-1} f(\hat{\boldsymbol{\theta}}_{SR}) \xrightarrow{d} \chi^2(1),$$

where  $\xrightarrow{d}$  denotes convergence in distribution.

This procedure can then be applied to the predictive regression Fama and French (1988) use to test for mean reversion

$$\tilde{r}_{t+q}(q) = \alpha_q + \beta_q \tilde{r}_t(q) + \varepsilon_{t+q}, \quad (23)$$

albeit, with the regression specified in terms of excess, rather than actual, returns.

This test requires us to condition on our estimated state variables. The state variables in the Bansal-Yaron model are  $x_t$  and  $\sigma_t^2$ . The sole state variable for the Campbell-Cochrane model is  $s_t$ . We extract  $x_t$ ,  $\sigma_t^2$  and  $s_t$  as explained earlier. The sole state variable for the Cecchetti-Lam-Mark model is  $y_t$ , which can be extracted by estimating the Markov-switching model for consumption in isolation and taking  $y_t = 1$  if  $\Pr(y_t = 1 | \mathcal{F}_{t+1}) \geq \frac{1}{2}$ , where  $\mathcal{F}_t$  is information available at  $t$ .

## 5 Data

Data for our main results are from the US from 1947Q1 - 2017Q1. The time period is quarterly and, as is standard in the asset pricing literature, the agent's decision interval is assumed to be the time horizon considered.<sup>17</sup> The market index is the value-weighted CRSP index, obtained from WRDS. The set of assets also includes the six double-sorted size/book-to-market portfolios from Ken French's website. The risk-free rate will be the US one-month Treasury bill, also from French's website. Consumption data will be seasonally adjusted per-capita non-durables and services personal consumption expenditures, taken from the BEA. We deflate nominal data by the BEA's consumption deflator. For robustness, we consider results using annual data. Descriptive statistics of the quarterly data are given in Table 1.

<sup>17</sup>Constantinides and Ghosh (2011) suggest this makes little difference for the Bansal-Yaron model.



Table 1: Data summary statistics

	Mean	Median	Std dev	SC(1)
$r_m$	0.018	0.029	0.081	0.077
$r_f$	0.002	0.003	0.007	0.745
$\Delta c$	0.005	0.006	0.005	0.279
$\Delta d$	0.007	0.001	0.148	0.584
$z_m$	4.871	4.851	0.426	0.937

$r_m$  denotes the market return,  $r_f$  the quarterly risk-free rate (the rolled over 1 month US T-bill),  $\Delta c$  log consumption growth,  $\Delta d$  log dividend growth and  $z_m$  the log price-dividend ratio.

## 5.1 Serial dependence in the market return

To provide context to the results regarding serial dependence in the model residuals, we present statistics regarding serial dependence in the market return below in Table 3. Table 3a shows the variance ratio statistics. We see that, for all but  $q = 8$ , the variance ratio point estimates are somewhat below one, suggesting mean reversion in the market return, although this is not significant. However, the standard errors are large. To reject at the 10% level given the standard error at  $q = 40$ , the variance ratio, which is equal to one under the null and bounded from below by 0, would have to be 0.28 or less. This underscores Poterba and Summers's (1988) point that the variance ratio has low power against a mean reverting alternative. Variance ratio estimates can be biased downwards when  $q/T$  is not small (Campbell et al., 1997). However, given the known low power of variance ratio tests against a mean reverting alternative and that the variance ratios from  $q = 16$  onwards are somewhat below one, where  $q/T$  is fairly small, there is at least some evidence of mean reversion.

This evidence becomes stronger when we look at sub-samples of the data. Figure 2 shows moving window estimates of the variance ratios for different holding periods. The variance ratios are calculated over samples of 30 years (120 quarters) and the window start date moves on by 10 years each iteration. With one exception, all of the point estimates are below 1, and in most cases substantially so. However, even though the variance ratio estimates become very low, dropping below 0.2, none of them is significantly different from one at conventional significance levels in one or two-tailed tests. Nonetheless, the estimates do provide at least indicative evidence of mean reversion in the market return.

Table 3b shows the estimated weighted quantilograms for the market for a variety of holding period-quantile combinations. Bootstrapped  $p$ -values are shown in smaller font underneath the point estimate. We calculate these by applying a standard normal wild bootstrap to market returns. The quantilograms show little evidence of directional predictability, both in terms of the point estimates and the  $p$ -values. We calculate the Hurst (1951) re-scaled range statistic in Table 3c. This provides no significant evidence of serial dependence in quarterly market returns, either.

## 5.2 Model estimation

Some of the Bansal-Yaron model parameter estimates Table 4 are surprising. Mean dividend growth is estimated to be negative and the long-run risk is non-stationary ( $\hat{\rho}_x > 1$ ). However, the standard errors of these parameters are fairly large so the point estimates

Table 3: Serial dependence in the market return

(a) Variance ratio

$q$	8	12	16	20	24	28	32	36	40
$VR$	0.963	0.857	0.795	0.833	0.809	0.797	0.795	0.807	0.801
(Std err)	(0.206)	(0.254)	(0.291)	(0.323)	(0.351)	(0.377)	(0.400)	(0.423)	(0.434)
$p$ -value	0.857	0.573	0.482	0.604	0.587	0.590	0.608	0.649	0.654

(b) Quantilogram

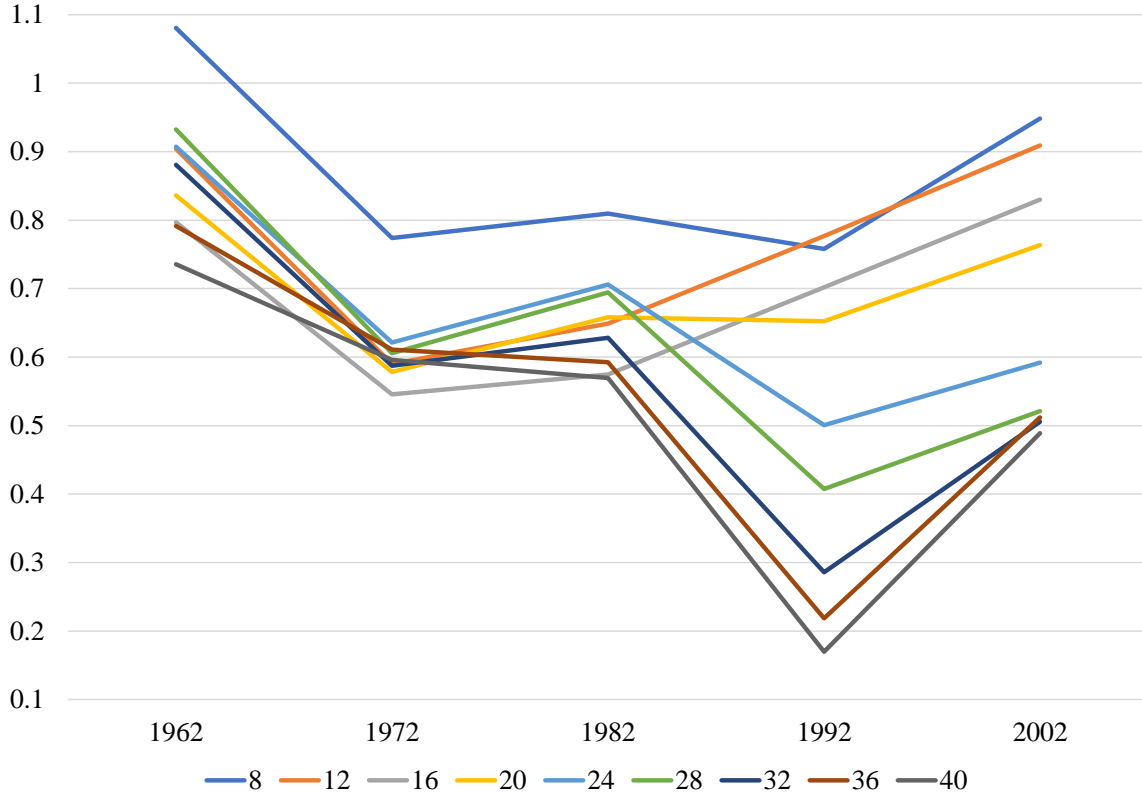
$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	-0.035	0.175	0.371	0.106	-0.012	-0.021	-0.050	0.100	0.032
	0.450	0.390	0.530	0.740	0.890	0.620	0.490	0.450	0.770
0.05	-0.055	0.175	0.435	0.053	-0.052	-0.031	-0.092	0.073	0.031
	0.610	0.510	0.530	0.680	0.840	0.620	0.490	0.450	0.980
0.1	-0.076	0.150	0.448	0.041	-0.065	-0.047	-0.129	0.043	0.017
	-0.720	0.500	0.520	0.640	0.790	0.600	0.450	0.420	0.860
0.25	-0.097	0.112	0.418	0.116	-0.022	-0.030	-0.167	0.034	-0.002
	0.910	0.430	0.520	0.570	0.730	0.550	0.440	0.390	0.650
0.5	-0.118	0.062	0.347	0.138	-0.008	-0.023	-0.190	0.028	-0.024
	0.900	0.430	0.540	0.560	0.640	0.550	0.370	0.370	0.600
0.75	-0.139	0.034	0.293	0.130	-0.020	-0.047	-0.203	0.038	-0.047
	0.780	0.430	0.590	0.570	0.620	0.520	0.410	0.360	0.530
0.9	-0.161	0.020	0.275	0.129	-0.022	-0.065	-0.221	0.036	-0.072
	0.570	0.450	0.600	0.580	0.570	0.510	0.420	0.360	0.450
0.95	-0.182	0.004	0.293	0.148	-0.009	-0.070	-0.225	0.042	-0.097
	0.440	0.410	0.570	0.570	0.500	0.500	0.420	0.370	0.420
0.99	-0.204	-0.007	0.317	0.169	-0.013	-0.081	-0.247	0.044	-0.123
	0.400	0.380	0.520	0.560	0.450	0.450	0.380	0.330	0.380

(c) Rescaled range

R/S	0.957
$p$ -value	0.258

$VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

Figure 2: Moving window variance ratio estimates



The variance ratios are calculated over 30 year windows centred on Q2 of the given year.

Table 4: Bansal-Yaron model estimates

$\mu_c$	$\mu_d$	$\phi$	$\varphi$	$\rho_x$	$\psi_x$	$\sigma$	$\nu$	$\sigma_w$	$\delta$	$\psi$	$\gamma$
0.004	-0.023	4.341	6.221	1.160	-0.085	0.008	0.458	$-9.5 \times 10^{-5}$	0.995	1.890	6.994
(14.11)	(1.740)	(0.006)	(0.005)	(0.275)	(0.026)	(2.447)	(0.045)	(0.011)	(0.454)	(0.058)	(0.004)
1.000	0.990	0	0	$2.5 \times 10^{-5}$	0.001	0.997	0	0.993	0	0	0.028
$J$ -stat	415.8	$p$ -value	0								

Point estimates are displayed in the first row, standard errors (in parentheses) in the second and  $p$ -values in the third. All  $p$ -values are asymptotic.

are all less than one standard error away from a “reasonable” value. However, the  $J$ -statistic  $p$ -value is negligible and the model specification is heavily rejected.

Nonetheless, the preference parameters are reasonable: the subjective discount factor is less than one and the risk-aversion coefficient significantly greater than zero. The population mean market return,  $E(r_{m,t})$ , is 0.35% on a quarterly basis.<sup>18</sup> This is a little on the low side, but still reasonable. However, the mean of our plug-in estimator of  $E_t r_{m,t+1}$  is -84% on a quarterly basis. This is clearly very unreasonable and arises because the mean of  $\sigma_t^2$  over  $t$  is 0.018, which is substantially larger than the theoretical mean,  $\sigma^2 = 7.1 \times 10^{-5}$ .

As can be seen in Table 5, the Campbell-Cochrane model parameter estimates are reasonable. The estimate of  $\bar{g}$  is close to both the sample mean of  $\Delta c$  and  $\Delta d$ , which are in any case similar, and the estimates of  $\text{Var}(\Delta c)$  and  $\text{Var}(\Delta d)$  are also both close to their sample counterparts. The preference parameter estimation shows that the subjective

<sup>18</sup> $E(r_{m,t}) = B_0 + B_2\sigma^2$  since  $E(x_t) = 0$  and  $E(\sigma_t^2) = \sigma^2$ .

Table 5: Campbell-Cochrane model estimates

$\bar{g}$	Var( $\Delta c$ )	Var( $\Delta d$ )	Cov( $\Delta c, \Delta d$ )	$\alpha$	$\phi$	$\delta$	$\gamma$
0.005 ( $4.73 \times 10^{-4}$ )	$2.83 \times 10^{-5}$ ( $4.04 \times 10^{-6}$ )	0.021 (0.004)	$-8.44 \times 10^{-6}$ ( $6.37 \times 10^{-5}$ )	0.316 (0.103)	0.936 (0.021)	0.41 (0.274)	22.29 (7.537)
0	0	0	0.133	0.002	0	0.135	0.003
$J$ -stat	0.387			$R^2$	0.877	$J$ -stat	8.837
$p$ -value	0.534					$p$ -value	0.116

Each panel (set of columns) refers to a separate estimation. The estimates of  $\delta$  and  $\gamma$  condition on the estimates in the first two panels. Point estimates are displayed in the first row, standard errors (in parentheses) in the second and  $p$ -values in the third. All  $p$ -values are asymptotic.

discount factor is also significantly less than 1 and the estimated utility curvature is significantly greater than zero. The  $J$ -test does not reject the model's Euler equations. Nonetheless, this is only indicative of model performance, since the hypothesis tests referred to in the preference parameter GMM condition on the earlier estimates of the model time-series parameters.

However, Table 7 shows that the average expected return (-37%) is extreme and a long way from the sample mean return (1.8%). Moreover, the estimated subjective discount factor is rather low at 0.41 from quarter to quarter. However, its standard error is large. The estimated utility curvature is also high at 22. This makes average risk aversion very high once interacted with the state variable. Assuming a Lucas tree-style endowment economy, average risk aversion in sample is 236.<sup>19</sup> Still, Campbell and Cochrane (1999) give good reasons to not be unduly concerned by high average risk-aversion. Despite these issues in estimating the Campbell-Cochrane model, especially the large mean pricing error, the standard formal statistical test of the model (the  $J$ -test) is unable to reject it. Hence we see the value of examining second moments.

Table 6 shows that the Cecchetti-Lam-Mark model is reasonably estimated, but not as well as the Campbell-Cochrane model. The preference parameters give rise to an insignificant  $J$ -statistic. However, the risk aversion coefficient is high at 240, although this is similar to the average risk aversion in the Campbell-Cochrane model. Moreover, the subjective discount factor is greater than 1, but has a large standard error. Our estimated Markov-switching model is somewhat different to that of Cecchetti et al. (1990). The two states are high and low growth states: both states have positive mean consumption growth. As a consequence, the unconditional probability of the bad (low-growth) state is much larger than in the original Cecchetti et al. (1990) paper. However, the implied unconditional mean of consumption growth is 0.5%. Expected returns are a shade low at 0.8% per quarter.

## 6 Serial dependence in the model residuals

Our results for the Bansal-Yaron model are shown in Table 8. There is very clear evidence that the model residuals are not a martingale difference sequence. The Bansal-Yaron model does not adequately capture the serial dependence shown in the market.

<sup>19</sup>Local utility curvature at  $t$  is given by  $\gamma/S_t$ . In an endowment economy where the only source of income and consumption is dividends, the elasticity of consumption with respect to wealth is 1, so local utility curvature is equal to relative risk aversion. C.f. Campbell and Cochrane (1999, p.p. 244).

Table 6: Cecchetti-Lam-Mark model estimates

(a) Consumption model

$\alpha_0$	$\alpha_1$	$p$	$q$	$\pi$	$\sigma_0^2$	$\sigma_1^2$
0.007	-0.004	0.939	0.916	0.423	0.003	0.007

(b) Preference parameters

$\delta$	$\gamma$
1.234	240
(0.402)	(80.5)
0.002	0.003
$J$ -stat	7.439
$p$ -value	0.190

Panel (a) presents point estimates only. In panel (b), point estimates are displayed in the first row, standard errors (in parentheses) in the second and  $p$ -values in the third. All  $p$ -values are asymptotic.

Table 7: Properties of  $\hat{\xi}_t$ 

Model	Mean	Median	Std dev	SC(1)
Bansal-Yaron	0.858	0.907	0.494	0.874
Campbell-Cochrane	-0.351	-0.341	0.091	0.264
Cecchetti-Lam-Mark	0.010	0.018	0.082	0.072

SC(1) denotes first-order serial correlation

The variance ratios are all significantly different from one. They are all also greater than one, suggesting that the residuals are positively serially correlated. The quantilogram tests provide overwhelming evidence of non-linear predictability in the model residuals. 71 of the 81 weighted quantilograms are significant, 70 of those at the 1% level. Only the 99th percentile shows no significant predictability. In almost all cases, except for the 1st percentile, the observed weighted quantilograms are greater than bootstrap means. This pattern reverses for the 1st percentile. The modified rescaled range test also overwhelmingly rejects the null that the residuals are MDS. The test gives rise to a negligible  $p$ -value. This time, the rejection is in the left tail of the distribution, such that the residuals show anti-persistent long-range dependence. Given that the mean residual shows that the Bansal-Yaron model is clearly cannot price the market, this finding is unsurprising.

More interestingly, the maximal predictability results suggest that the Bansal-Yaron state variables do not provide a convincing explanation of observed predictability, either. There are extremely significant exceedences of the  $R^2$  bound at six horizons: 8, 12, 20, 24, 32 and 36 quarters. However, we express some caution regarding these results for two reasons. First, the  $R^2$  bound is, for the Bansal-Yaron model, almost always either less than zero or greater than one for the holding periods considered. So either any degree of predictability is consistent with the long-run risk and time-varying economic volatility being risk factors in the stochastic discount factor or no predictability is consistent with these risk factors. Second, the  $R^2$  and  $\bar{R}^2$  reported are computed using plug-in estimators of moments from GMM, not from regressions themselves. The methods ought to be equivalent but it is not computationally possible to satisfy the moment conditions exactly

Table 8: Bansal-Yaron serial dependence results

(a) Variance ratios

$q$	8	12	16	20	24	28	32	36	40
$VR(q)$	<b>1.453</b>	<b>2.965</b>	<b>5.614</b>	<b>8.152</b>	<b>10.13</b>	<b>9.561</b>	<b>10.56</b>	<b>11.09</b>	<b>10.33</b>
(Std Err)	(0.217)	(0.262)	(0.311)	(0.362)	(0.408)	(0.465)	(0.495)	(0.524)	(0.562)
$p$ -value	0.037	0	0	0	0	0	0	0	0

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	<b>0.170</b>	<b>0.167</b>	<b>0.300</b>	<b>0.401</b>	<b>0.469</b>	<b>0.517</b>	<b>0.554</b>	<b>0.582</b>	0.605
	0	0	0	0	0	0	0	0.06	0.13
0.05	<b>1.322</b>	<b>1.750</b>	<b>2.196</b>	<b>2.549</b>	<b>2.786</b>	<b>2.948</b>	<b>3.062</b>	<b>3.145</b>	<b>3.204</b>
	0	0	0	0	0	0	0	0	0
0.1	<b>1.347</b>	<b>1.896</b>	<b>2.478</b>	<b>2.936</b>	<b>3.277</b>	<b>3.531</b>	<b>3.716</b>	<b>3.844</b>	<b>3.929</b>
	0	0	0	0	0	0	0	0	0
0.25	<b>2.526</b>	<b>3.697</b>	<b>4.762</b>	<b>5.705</b>	<b>6.523</b>	<b>7.193</b>	<b>7.701</b>	<b>8.075</b>	<b>8.349</b>
	0	0	0	0	0	0	0	0	0
0.5	<b>1.868</b>	<b>2.825</b>	<b>3.709</b>	<b>4.509</b>	<b>5.233</b>	<b>5.877</b>	<b>6.465</b>	<b>6.997</b>	<b>7.473</b>
	0	0	0	0	0	0	0	0	0
0.75	<b>1.970</b>	<b>2.935</b>	<b>3.801</b>	<b>4.620</b>	<b>5.408</b>	<b>6.127</b>	<b>6.765</b>	<b>7.331</b>	<b>7.852</b>
	0	0	0	0	0	0	0	0	0
0.9	<b>2.468</b>	<b>3.623</b>	<b>4.577</b>	<b>5.342</b>	<b>5.970</b>	<b>6.495</b>	<b>6.915</b>	<b>7.254</b>	<b>7.533</b>
	0	0	0	0	0	0	0	0	0
0.95	<b>2.226</b>	<b>3.164</b>	<b>3.891</b>	<b>4.361</b>	<b>4.638</b>	<b>4.801</b>	<b>4.893</b>	<b>4.937</b>	<b>4.947</b>
	0.01	0	0	0	0	0	0	0	0
0.99	0.480	0.719	0.845	0.910	0.944	0.960	0.965	0.963	0.955
	0.96	0.57	0.9	0.97	0.89	0.88	0.85	0.83	0.86

(c) Rescaled range

R/S	<b>10.08</b>
$p$ -value	0

(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
$R^2$	0.057	0.018	$10^{-8}$	0.069	0.025	0.012	0.119	0.011	0.159
$\bar{R}^2$	-0.217	-11.357	5.508	-3.740	-0.282	0.303	-10.089	-6.736	6.579
Wald stat	<b>216041</b>	<b>148018</b>	-	<b>267587</b>	<b>416715</b>	-	<b>593143</b>	<b>1307207</b>	-
$p$ -value	0	0	-	0	0	-	0	0	-

Bold values denote significance at the 10% level against the appropriate null hypothesis. For all except panel (d), this is the martingale difference hypothesis. For panel (d), the null is that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

here, despite the model being exactly identified. Therefore the methods are not equivalent in a finite sample. Because of this, the reported  $R^2$  for the predictive regression for a given horizon is not the same for each model, even though it should be. Comparison of Table 8 with Tables 9 and 10 shows that the reported sample  $R^2$  for a given  $q$  is not always particularly similar. This highlights that the numerical challenges of the GMM estimation undertaken to compute the tests.

Our main results regarding the Campbell-Cochrane model are in Table 9. The variance ratios appear to show that the Campbell-Cochrane model of returns cannot explain the serial dependence exhibited by the market return. The variance ratios are predominantly greater than one, and those which are significant<sup>20</sup> are all greater than one. There is thus evidence of positive serial correlation in the residuals.

Those variance ratios which are significant are generally substantially greater than one, suggesting that the positive serial correlation in the Campbell-Cochrane residuals is substantial. Moreover, the significance of the variance ratios comes at a variety of horizons. It is not concentrated over short or long holding periods.

The quantilogram tests provide further evidence that the Campbell-Cochrane model cannot explain the serial dependence structure of the market return. 35 of the 81 weighted quantilograms are significant. The clear evidence of quantile predictability in the higher quantiles puts the evidence from the variance ratios, that the Campbell-Cochrane residuals are not MDS, beyond doubt. In all of the violations, the sample weighted quantilogram is less than the mean of the bootstrap distribution, suggesting that the sample quantilograms are too small. The mean of all of the bootstrap distributions is positive, suggesting that, under the null, the Campbell-Cochrane residuals ought to display some positive quantile predictability. The apparent lack of quantile predictability in the market return suggests that the presence of predictability under the null is a finite-sample phenomenon, induced by the estimation of the model parameters. The rescaled range statistic buttresses the evidence of the variance ratios and the quantilogram, in that it is highly significant. Again, the rejection comes in the right tail of the rescaled range's distribution.

Having seen that the Campbell-Cochrane model cannot adequately capture the degree of own-history predictability in the market return, we consider whether the surplus consumption state variable  $S_t$  is able to. We use our modified Huang and Zhou (2015) Wald test. There are only two exceedences of the  $R^2$  bound in regressions of  $q$ -period returns  $r_{t+q}(q)$  on the lagged  $q$ -period return  $r_t(q)$ . These are both highly significant and occur at the two shortest holding periods considered. The lack of exceedence at longer holding periods is due to a mix of low  $R^2$ s and loose bounds. Untabulated results confirm that violations bound seem to be concentrated at shorter holding periods. The  $R^2$  bound is exceeded for  $q = 3, 4, 5$  and 6 quarters these exceedences are significant at the 1% level. The evidence, then, is that the  $S_t$  state variable has difficulty explaining the degree of predictability seen in the market, which is often too great.

Table 10 shows the results for the Cecchetti-Lam-Mark model. Few of the test statistics are significant, consistent with Cecchetti et al.'s (1990) original finding. However, the variance ratios from  $q = 24$  onwards are rather different from one. At around 0.4, the standard errors on the variance ratios are large from  $q = 24$  quarters onwards. With the variance ratios becoming as low as 0.62, there does appear to be some negative serial correlation in the residuals. Although variance ratios can be biased downwards when  $q/T$  is not small, the variance ratios are substantially below one even for  $q/T$  fairly small (e.g.

---

<sup>20</sup>Recall that we use the 10% level due to the arguments of Poterba and Summers (1988).

Table 9: Campbell-Cochrane serial dependence results  
(a) Variance ratios

$q$	8	12	16	20	24	28	32	36	40
$VR$	<b>1.492</b>	1.346	<b>3.423</b>	0.928	1.423	<b>2.541</b>	0.435	1.602	<b>2.541</b>
(Std err)	(0.196)	(0.253)	(0.278)	(0.331)	(0.358)	(0.382)	(0.416)	(0.442)	(0.463)
$p$ -value	0.012	0.171	0	0.830	0.237	$5 \times 10^{-5}$	0.175	0.173	$9 \times 10^{-4}$

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	-0.035	-0.037	-0.006	0.004	0.003	-0.003	-0.012	-0.025	-0.038
	0.63	0.69	0.72	0.76	0.8	0.8	0.83	0.84	0.87
0.05	0.124	0.126	0.134	0.155	0.146	0.146	0.175	0.219	0.261
	0.85	0.99	0.95	0.97	0.95	0.93	0.96	1	0.98
0.1	0.522	0.761	0.987	1.188	1.344	1.512	1.699	1.871	2.016
	0.36	0.4	0.4	0.41	0.43	0.39	0.4	0.35	0.3
0.25	1.023	1.561	2.041	2.469	2.856	3.229	3.582	3.914	4.203
	0.58	0.58	0.58	0.58	0.58	0.58	0.59	0.59	0.59
0.5	0.392	0.637	0.872	1.135	1.352	1.548	1.722	1.906	2.080
	0.68	0.68	0.69	0.75	0.8	0.81	0.83	0.85	0.85
0.75	<b>0.252</b>	<b>0.415</b>	<b>0.566</b>	<b>0.726</b>	<b>0.862</b>	<b>0.966</b>	<b>1.043</b>	<b>1.127</b>	<b>1.184</b>
	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
0.9	<b>0.151</b>	<b>0.164</b>	<b>0.190</b>	<b>0.256</b>	<b>0.330</b>	<b>0.376</b>	<b>0.413</b>	<b>0.473</b>	<b>0.512</b>
	0.04	0.03	0.02	0.02	0.03	0.03	0.02	0.02	0.02
0.95	<b>0.146</b>	<b>0.146</b>	<b>0.148</b>	<b>0.169</b>	<b>0.227</b>	<b>0.266</b>	<b>0.306</b>	<b>0.356</b>	<b>0.391</b>
	0.08	0.05	0.04	0.04	0.03	0.04	0.05	0.05	0.05
0.99	<b>-0.046</b>	<b>-0.073</b>	<b>-0.080</b>	<b>-0.049</b>	<b>-0.011</b>	<b>0.008</b>	<b>0.015</b>	<b>0.032</b>	0.040
	0.08	0.05	0.03	0.04	0.07	0.08	0.08	0.08	0.21

(c) Rescaled range

R/S	<b>4.964</b>
$p$ -value	0

(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
$R^2$	0.029	0.041	0.059	$1.2 \times 10^{-4}$	0.001	0.005	$1.4 \times 10^{-4}$	0.006	0.082
$\bar{R}^2$	0.003	0.018	0.421	0.039	0.006	0.556	0.352	1.505	0.689
Wald stat	<b>462.1</b>	<b>7.954</b>	-	-	-	-	-	-	-
$p$ -value	0	0.005	-	-	-	-	-	-	-

Bold values denote significance at the 10% level against the appropriate null hypothesis. For all except panel (d), this is the martingale difference hypothesis. For panel (d), the null is that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.



Table 10: Cecchetti-Lam-Mark model results  
(a) Variance ratio

$q$	8	12	16	20	24	28	32	36	40
$VR$	0.910	0.830	0.757	0.733	0.690	0.646	0.621	0.645	0.643
(Std err)	(0.213)	(0.268)	(0.306)	(0.337)	(0.366)	(0.394)	(0.422)	(0.445)	(0.468)
$p$ -value	0.671	0.527	0.427	0.427	0.391	0.369	0.369	0.425	0.445

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	-0.035	0.124	0.324	0.119	-0.015	0.010	-0.036	0.099	0.032
	0.28	0.44	0.03	0.32	0.71	0.33	0.15	0.34	0.66
0.05	-0.056	0.127	0.361	0.083	-0.057	-0.017	-0.067	0.072	0.031
	0.46	0.6	0.04	0.61	0.88	0.29	0.17	0.42	0.66
0.1	-0.076	0.139	0.358	0.083	-0.067	-0.048	-0.123	0.037	0.017
	0.54	0.6	0.1	0.59	0.86	0.24	0.11	0.51	0.69
0.25	-0.097	0.142	0.326	0.157	-0.026	-0.034	<b>-0.178</b>	0.021	-0.002
	0.72	0.6	0.16	0.24	0.59	0.35	0.06	0.49	0.78
0.5	-0.118	0.120	0.255	0.175	-0.015	-0.024	<b>-0.212</b>	0.011	-0.024
	0.74	0.69	0.35	0.2	0.53	0.55	0.06	0.45	0.88
0.75	-0.140	0.100	0.202	0.170	-0.031	-0.041	<b>-0.233</b>	0.017	-0.048
	0.97	0.74	0.49	0.19	0.53	0.62	0.05	0.33	0.9
0.9	-0.161	0.082	0.195	0.170	-0.034	-0.048	<b>-0.251</b>	0.012	-0.072
	0.87	0.72	0.48	0.18	0.51	0.75	0.07	0.3	0.97
0.95	-0.183	0.071	0.223	0.187	-0.020	-0.033	<b>-0.265</b>	0.007	-0.098
	0.67	0.73	0.38	0.15	0.46	0.92	0.08	0.3	0.9
0.99	-0.205	0.059	0.253	0.204	-0.022	-0.031	<b>-0.291</b>	-0.002	-0.124
	0.65	0.78	0.32	0.12	0.42	0.96	0.06	0.3	0.78

(c) Rescaled range

R/S	1.653
$p$ -value	0.168

(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
$R^2$	0.367	0.011	0.105	0.005	0.001	0.099	0.012	0.023	0.072
$\bar{R}^2$	1.298	0.347	0.411	3.858	$5.6 \times 10^{-7}$	0.499	3.071	3.196	0.861
Wald stat	-	-	-	-	<b>10.95</b>	-	-	-	-
$p$ -value	-	-	-	-	$9.4 \times 10^{-4}$	-	-	-	-

Bold values denote significance at the 10% level against the appropriate null hypothesis. For all except panel (d), this is the martingale difference hypothesis. For panel (d), the null is that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

at  $VR(24) = 0.68$  and  $q/T = 0.09$ ). Bias alone seems unlikely to be enough to explain these findings. The evidence that the Cecchetti-Lam-Mark model residuals are MDS is certainly unconvincing.

The weighted quantilegram results show little evidence of non-linear predictability in the Cecchetti-Lam-Mark residuals. The point estimates are, by-and-large, close to zero. Only at the 32 quarter holding period are any of the results significant. The rescaled range statistic is also insignificant with a  $p$ -value of 0.17.

Considering the modified Huang and Zhou (2015) Wald test, we see that, for longer horizons (i.e. two or more years), the consumption regime state variable, indicating a high or a low growth state, does a reasonable job of explaining the degree of predictability observed. There is only one exceedence of the  $R^2$  bound, at  $q = 24$  quarters, although this is very significant.

## 6.1 Sub-sample analysis

One of the main criticisms of evidence of mean reversion has been its apparent sensitivity to the time period chosen (e.g. Kim et al., 1991). We see in Figure 2 that the strength of the evidence of mean reversion does indeed appear to change over time. We therefore investigate the sensitivity of our results to the time period used by repeating our analysis over two sub-samples. The first runs from 1947Q1-1982Q1 and the second 1982Q2-2017Q1. Dividing the sample into two ensures a sample size in excess of 120 (i.e.  $3 \times \max\{q\}$ ) in each sub-sample, which is necessary for the estimation of the long-horizon serial correlations.

While the form of departure from an MDS appears to change over the sub-samples for the Bansal-Yaron and Campbell-Cochrane models, there are clear departures from the MDS null in both sub-samples for these models. The rescaled range results are robust across the sub-samples for the Bansal-Yaron and Campbell-Cochrane models. However, it produces a rejection in the first sub-sample for the Cecchetti-Lam-Mark model, which it does not in the second. The variance ratio and quantilegram findings are robust for the Cecchetti-Lam-Mark model. Only for the Bansal-Yaron model are the maximal predictability results broadly similar across the sub-samples.

Table 11 shows the sub-sample results for the Bansal-Yaron model. We see that the variance ratio results are very different across the sub-samples. The results for the second sub-sample are similar to those for the sample as a whole: strong evidence of positive autocorrelation in the residuals with vanishingly small  $p$ -values. However, in the first sub-sample, we see insignificant evidence of negative serial correlation. The estimated variance ratios are again rather different from one, seemingly highlighting the low power of variance-ratio tests against mean reverting alternatives. In fact, from  $q = 32$  onwards, the standard errors are sufficiently large such that the martingale difference hypothesis would never be rejected at the 10% level if the estimated variance ratio were less than one. Given how different the variance ratios are from one, the residuals do not seem to be MDS in the first sub-sample, either. It appears the positive autocorrelation found in the second sub-sample dominates in the whole sample.

The quantilegram evidence shows some agreement between the two sub-samples, although the evidence of predictability is stronger in the first sub-period than in the second. Only for the 90th percentile are there large differences between the two sub-samples. The two sub-samples are also mainly in line with the whole sample, where the evidence of predictability is much stronger. The significant  $p$ -values from the sub-samples are closer

Table 11: Bansal-Yaron model sub-sample results  
(a) Variance ratios

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$VR(q)$	0.762	0.555	0.450	0.483	0.545	0.490	0.418	0.401	0.365
(Std err)	(0.341)	(0.413)	(0.466)	(0.512)	(0.555)	(0.596)	(0.633)	(0.668)	(0.700)
$p$ -value	0.484	0.282	0.238	0.313	0.412	0.392	0.358	0.370	0.364
Sub-sample 2: 1982Q2-2017Q1									
$VR(q)$	<b>4.485</b>	<b>6.931</b>	<b>8.870</b>	<b>10.21</b>	<b>11.66</b>	<b>12.89</b>	<b>13.73</b>	<b>13.72</b>	<b>11.84</b>
(Std err)	(0.354)	(0.486)	(0.565)	(0.630)	(0.668)	(0.699)	(0.728)	(0.755)	(0.767)
$p$ -value	0	0	0	0	0	0	0	0	0

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	<b>0.052</b>	<b>0.100</b>	<b>0.154</b>	<b>0.210</b>	<b>0.265</b>	<b>0.320</b>	<b>0.373</b>	<b>0.424</b>	<b>0.473</b>
	<b>0.050</b>	<b>0.075</b>	0.099	0.120	0.139	0.156	0.188	0.405	0.784
0.05	<b>0.567</b>	0.813	0.953	1.069	1.138	1.181	1.208	1.222	1.227
	0.115	0.120	0.082	0.026	-0.040	-0.060	-0.025	-0.001	-0.003
0.1	<b>0.987</b>	<b>1.291</b>	<b>1.579</b>	<b>1.896</b>	<b>2.127</b>	<b>2.268</b>	<b>2.358</b>	<b>2.413</b>	<b>2.442</b>
	0.493	<b>0.527</b>	<b>0.472</b>	<b>0.421</b>	<b>0.370</b>	<b>0.334</b>	<b>0.337</b>	<b>0.356</b>	<b>0.351</b>
0.25	<b>1.184</b>	<b>1.773</b>	<b>2.360</b>	<b>2.890</b>	<b>3.311</b>	<b>3.631</b>	<b>3.882</b>	<b>4.067</b>	<b>4.188</b>
	<b>1.001</b>	<b>1.131</b>	<b>1.067</b>	<b>0.933</b>	<b>0.753</b>	<b>0.595</b>	<b>0.485</b>	<b>0.419</b>	<b>0.349</b>
0.5	<b>1.621</b>	<b>2.433</b>	<b>3.203</b>	<b>3.916</b>	<b>4.503</b>	<b>4.967</b>	<b>5.307</b>	<b>5.557</b>	<b>5.718</b>
	<b>0.492</b>	<b>0.603</b>	<b>0.637</b>	<b>0.672</b>	0.654	0.586	0.486	0.420	0.369
0.75	<b>1.227</b>	<b>1.729</b>	<b>2.192</b>	<b>2.647</b>	<b>3.009</b>	<b>3.288</b>	<b>3.516</b>	<b>3.685</b>	<b>3.773</b>
	0.108	<b>0.165</b>	<b>0.189</b>	<b>0.206</b>	<b>0.191</b>	<b>0.139</b>	<b>0.057</b>	<b>-0.009</b>	-0.070
0.9	<b>0.878</b>	<b>0.983</b>	<b>1.064</b>	<b>1.245</b>	<b>1.410</b>	<b>1.520</b>	<b>1.577</b>	<b>1.604</b>	<b>1.599</b>
	0.023	0.022	-0.014	-0.069	-0.105	-0.145	-0.175	-0.173	-0.204
0.95	0.412	0.453	0.461	0.539	0.638	0.723	0.779	0.791	0.773
	-0.109	-0.116	-0.131	-0.146	-0.164	-0.206	-0.264	-0.320	-0.380
0.99	<b>0.207</b>	<b>0.230</b>	<b>0.228</b>	<b>0.215</b>	<b>0.197</b>	<b>0.175</b>	<b>0.151</b>	<b>0.125</b>	<b>0.063</b>
	-0.047	-0.074	-0.101	-0.127	-0.148	-0.168	<b>-0.186</b>	<b>-0.203</b>	<b>-0.220</b>

(c) Rescaled range

	1947Q1-1982Q1	1982Q2-2017Q1
R/S	<b>5.951</b>	<b>5.951</b>
$p$ -value	0	0

Bold values denote significance at the 10% level against the martingale difference hypothesis.  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

Table 11: Bansal-Yaron model sub-sample results - continued  
(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$R^2$	0.040	0.076	0.359	0.175	0.044	0.125	0.464	0.241	2.060
$\bar{R}^2$	-1.530	-5.110	-2.578	-8.822	-0.307	4.894	-0.731	-0.710	-111.3
Wald stat	<b>28569</b>	<b>83507</b>	<b>161393</b>	<b>189610</b>	<b>288991</b>	-	<b>1439362</b>	<b>419193</b>	<b>10436</b>
$p$ -value	0	0	0	0	0	-	0	0	0
Sub-sample 2: 1982Q2-2017Q1									
$R^2$	0.039	0.064	0.327	0.443	0.036	0.107	1.378	1.070	2.161
$\bar{R}^2$	-6.869	-11.03	10.31	-35.43	3.074	-17.65	20.38	4.249	-5423
Wald stat	<b>48313</b>	<b>95011</b>	-	<b>220739</b>	-	<b>140643</b>	-	-	<b>2761</b>
$p$ -value	0	0	-	0	-	0	-	-	0

Bold values denote significance at the 10% level against the null that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $p$ -values are asymptotic.

to 0.1 than those from the whole sample. The evidence from the first sub-sample is more similar to the sample as a whole. While the quantilogram evidence against the Bansal-Yaron model is stronger in the first period, there is still relatively strong evidence against the model in the second. There are 33 rejections of the null of no predictability in the Bansal-Yaron residuals in 81 tests. Thus we see relatively robust evidence that the Bansal-Yaron residuals are predictable in the quantilograms, although this evidence was weaker in the more recent period. Nonetheless, it is in this period that the variance ratio evidence is stronger. So there remains clear evidence of predictability in both sub-samples. While the form of the MDS violation is different in both sub-samples, we are able to detect both by using a variety of test statistics, each with different power of different alternatives. Furthermore, the re-scaled range statistics are almost identical across the two sub-samples and both provide extremely strong evidence of positive serial dependence in the model residuals.

The maximal predictability results are broadly consistent across the sub-samples, albeit with there being slightly stronger evidence against the Bansal-Yaron state variables in the first sub-sample. Of the six  $R^2$  bound violations across the whole sample, three are accompanied by  $R^2$  bound exceedences in both sub-samples. There is one holding period for which there are  $R^2$  bound violations in both sub-samples but not in the whole sample. Thus the results are a little unstable. Again, the  $\bar{R}^2$ s are generally either greater than one and less than zero. Given the volume and significance of violations in both sub-samples, the overall picture of the evidence is that the Bansal-Yaron state variables struggle to explain the degree of predictability observed.

For the Campbell-Cochrane model, sub-sample results are given in Table 12. For the variance ratios, there are clear differences between the sub-samples. The results for the first sub-sample are very different to the whole sample. They are not significant, although some are rather different from one (the variance ratios range from 0.17 to 1.41). Several of the standard errors are greater than 0.5 and so very large. In the second period, the variance ratios are mostly significant and greater than one, similar to the whole sample.

The quantilogram confirms that the Campbell-Cochrane residuals in the first period

Table 12: Campbell-Cochrane model sub-sample results  
(a) Variance ratios

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$VR(q)$	1.303	0.934	0.571	1.111	1.415	1.135	0.167	0.356	0.449
(Std err)	(0.287)	(0.371)	(0.423)	(0.478)	(0.520)	(0.558)	(0.619)	(0.661)	(0.694)
$p$ -value	0.290	0.858	0.310	0.816	0.425	0.810	0.178	0.329	0.428
Sub-sample 2: 1982Q2-2017Q1									
$VR(q)$	<b>4.176</b>	<b>1.663</b>	<b>3.683</b>	1.252	1.642	<b>1.966</b>	<b>2.226</b>	<b>4.419</b>	0.347
(Std err)	(0.301)	(0.364)	(0.398)	(0.477)	(0.529)	(0.565)	(0.623)	(0.645)	(0.754)
$p$ -value	0	0.068	$2 \times 10^{-11}$	0.597	0.225	0.088	0.049	$1 \times 10^{-7}$	0.387

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	<b>0.050</b>	<b>0.354</b>	<b>0.469</b>	<b>0.975</b>	<b>1.401</b>	<b>1.228</b>	<b>0.766</b>	<b>0.155</b>	<b>-0.047</b>
	<b>0.050</b>	0.075	0.099	0.120	0.139	0.156	0.172	0.228	0.466
0.05	0.075	0.364	0.415	<b>1.106</b>	<b>1.732</b>	<b>1.534</b>	<b>1.069</b>	0.130	0.028
	0.207	0.272	0.310	0.327	0.322	0.373	0.466	0.548	0.672
0.1	<b>0.099</b>	0.312	0.319	1.105	<b>1.910</b>	<b>1.668</b>	<b>1.278</b>	0.137	0.076
	0.352	0.413	0.532	0.622	0.656	0.771	0.943	1.066	1.171
0.25	<b>0.120</b>	<b>0.240</b>	<b>0.232</b>	1.059	1.952	1.685	1.423	<b>0.113</b>	0.094
	0.742	1.000	1.202	1.361	1.502	1.712	1.913	2.035	2.063
0.5	<b>0.139</b>	0.166	0.146	0.949	<b>1.864</b>	<b>1.620</b>	<b>1.450</b>	0.070	0.096
	<b>1.888</b>	<b>2.683</b>	<b>3.244</b>	<b>3.628</b>	<b>3.905</b>	<b>4.110</b>	4.224	4.220	4.094
0.75	0.156	0.128	0.087	0.832	<b>1.735</b>	<b>1.509</b>	<b>1.431</b>	0.019	0.088
	2.044	2.958	3.698	4.292	4.753	5.100	5.326	5.438	5.459
0.9	0.188	0.183	0.086	0.723	<b>1.557</b>	<b>1.332</b>	<b>1.377</b>	-0.002	0.075
	1.121	1.674	2.143	2.484	2.718	2.881	2.983	3.041	3.068
0.95	<b>0.405</b>	0.249	0.082	<b>0.629</b>	<b>1.358</b>	<b>1.152</b>	<b>1.300</b>	-0.041	0.058
	<b>0.887</b>	1.134	1.232	1.267	1.275	1.267	1.248	1.221	1.189
0.99	<b>0.784</b>	<b>0.259</b>	0.088	<b>0.571</b>	<b>1.191</b>	<b>0.995</b>	<b>1.205</b>	-0.082	<b>0.038</b>
	<b>0.308</b>	<b>0.298</b>	<b>0.280</b>	<b>0.267</b>	<b>0.257</b>	<b>0.248</b>	<b>0.239</b>	<b>0.231</b>	<b>0.223</b>

(c) Rescaled range

	1947Q1-1982Q1	1982Q2-2017Q1
R/S	<b>3.121</b>	<b>8.677</b>
$p$ -value	$5 \times 10^{-7}$	0

Bold values denote significance at the 10% level against the martingale difference hypothesis.  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

Table 12: Campbell-Cochrane model sub-sample results  
(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$R^2$	1.516	0.084	0.256	0.357	0.203	0.110	2.151	0.376	0.351
$\bar{R}^2$	4.145	0.208	0.007	0.121	0.025	0.000	2.215	0.008	0.485
Wald stat	-	-	<b>426.2</b>	<b>608.4</b>	<b>75.39</b>	<b>403.2</b>	-	<b>570.9</b>	-
$p$ -value	-	-	0	0	0	0	-	0	-
Sub-sample 2: 1982Q2-2017Q1									
$R^2$	0.151	0.064	0.301	0.383	0.112	0.409	32.09	0.724	-0.004
$\bar{R}^2$	0.108	32.29	0.272	0.006	0.024	0.375	8.988	0.005	-0.172
Wald stat	<b>6.887</b>	-	1.076	<b>713.5</b>	<b>36.00</b>	<b>6.546</b>	<b>2004</b>	<b>4049</b>	<b>24.86</b>
$p$ -value	0.009	-	0.300	0	0	0.011	0	0	$6.15 \times 10^{-7}$

Bold values denote significance at the 10% level against the null that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $p$ -values are asymptotic.

are not MDS. 43 of the weighted quantilograms are significant, in many cases at the 1% level. Inversely to the variance-ratio evidence, many of the first sub-period quantilograms are largely insignificant. Thus, the quantilogram evidence, like the variance-ratio evidence, is unstable. But, where one does not show significant predictability in the residuals, the other does. The conclusion is the same. The Campbell-Cochrane model cannot explain the serial dependence of the market return. This again shows the value of using multiple test statistics to assess the null that the residuals are MDS. How the serial correlation structure of the Campbell-Cochrane expected returns departs from that of the market appears to vary over time. The rescaled range shows similar results in both sub-samples and provides highly significant evidence of serial dependence in the residuals both times.

The results of the maximal predictability test also differ across the two sub-samples. First, we see that there are some periods where there is a significant violation of the  $R^2$  bound in one sub-sample and not the other. Moreover, the two significant violations across the whole sample occur for  $q = 8$  and  $q = 12$ . Yet, the only significant violation for these holding periods in the sub-samples is in the first period for  $q = 8$ . There are also several holding periods ( $q = 20, 24, 28, 36$ ) where  $\bar{R}^2$  is significantly exceeded in both sub-samples but not across the whole period. It is not entirely clear why this is, and, as mentioned above, the difficulties of implementing the GMM estimation required for this test mean that we must take these results with some caution. Nonetheless, the sub-samples provide stronger evidence that the surplus consumption ratio state variable is unable to explain the level of predictability observed in market returns.

Table 13 presents the sub-sample analysis for the Cecchetti-Lam-Mark model. The variance ratio results are very much similar to those across the whole sample. There is no significant evidence against the null that  $\xi_t$  is a martingale difference sequence. However, the variance ratio estimates in each sub-sample are lower than those in the whole sample. There are no rejections because the standard errors become larger too. In fact, some of the standard errors are very large. The variance ratio estimates are as low as 0.35, which is substantially different from 1. However, the standard errors are such that a two-tailed

Table 13: Cecchetti-Lam-Mark model sub-sample results  
(a) Variance ratios

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$VR$	0.842	0.763	0.645	0.624	0.673	0.616	0.508	0.483	0.405
(Std err)	(0.289)	(0.380)	(0.474)	(0.528)	(0.570)	(0.617)	(0.657)	(0.686)	(0.704)
$p$ -value	0.585	0.532	0.454	0.477	0.566	0.533	0.454	0.452	0.397
Sub-sample 2: 1982Q2-2017Q1									
$VR$	0.826	0.828	0.743	0.714	0.617	0.542	0.535	0.461	0.349
(Std err)	(0.304)	(0.398)	(0.454)	(0.496)	(0.586)	(0.619)	(0.657)	(0.729)	(0.766)
$p$ -value	0.568	0.666	0.572	0.564	0.513	0.459	0.479	0.459	0.395

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	0.050	0.277	0.354	0.148	0.026	0.007	-0.129	0.083	-0.047
	<b>0.050</b>	-0.014	0.404	0.133	-0.147	0.021	0.083	-0.017	-0.047
0.05	0.075	0.299	<b>0.420</b>	0.028	-0.008	-0.065	-0.217	0.010	-0.074
	<b>0.075</b>	-0.038	0.487	0.176	-0.193	0.047	0.074	-0.037	-0.074
0.1	0.099	0.267	0.426	-0.012	-0.017	-0.099	<b>-0.314</b>	-0.069	-0.101
	<b>0.099</b>	-0.042	0.525	0.195	-0.206	0.038	0.048	-0.066	-0.101
0.25	0.120	0.281	0.426	0.067	0.045	-0.060	<b>-0.345</b>	-0.117	-0.089
	<b>0.120</b>	-0.091	0.485	0.224	-0.187	0.032	0.037	-0.078	-0.129
0.5	0.139	0.283	0.392	0.122	0.107	-0.011	<b>-0.365</b>	-0.124	-0.056
	0.142	-0.163	0.385	0.172	-0.208	-0.016	0.037	-0.103	-0.155
0.75	0.156	0.284	0.351	0.104	0.107	-0.015	<b>-0.397</b>	-0.112	-0.042
	0.171	-0.223	0.310	0.139	-0.230	-0.067	0.053	-0.102	-0.180
0.9	0.188	0.311	0.347	0.076	0.091	-0.017	<b>-0.418</b>	-0.134	-0.043
	0.206	-0.267	0.275	0.130	-0.234	-0.115	0.043	-0.126	-0.205
0.95	0.405	0.366	0.392	0.094	0.097	0.001	<b>-0.412</b>	-0.130	-0.060
	0.243	-0.311	0.277	0.126	-0.228	-0.145	0.019	-0.169	-0.229
0.99	0.784	0.378	0.469	0.138	0.115	0.018	<b>-0.393</b>	-0.106	-0.089
	0.280	-0.340	0.285	0.112	-0.252	-0.175	-0.026	-0.215	-0.253

(c) Rescaled range

	1947Q1-1982Q1	1982Q2-2017Q1
R/S	<b>2.512</b>	1.452
$p$ -value	$3 \times 10^{-4}$	0.439

Bold values denote significance at the 10% level against the martingale difference hypothesis.  $VR$  denotes the estimated variance ratio. The estimated weighted quantilogram is given in larger font for the appropriate  $(\alpha, q)$  combination. Its bootstrapped  $p$ -value is given underneath in smaller font. All other  $p$ -values are asymptotic.

Table 13: Cecchetti-Lam-Mark model sub-sample results - continued  
(d) Maximal predictability

$q$	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
$R^2$	0.072	0.072	0.137	0.228	0.106	0.153	0.276	0.630	0.416
$\bar{R}^2$	0.467	0.107	1.658	0.666	0.009	0.044	0.136	14.81	2.064
Wald stat	-	-	-	-	<b>494.3</b>	<b>806.1</b>	<b>553.0</b>	-	-
$p$ -value	-	-	-	-	0	0	0	-	-
Sub-sample 2: 1982Q2-2017Q1									
$R^2$	0.023	0.156	0.262	0.161	0.071	0.076	0.546	0.792	0.519
$\bar{R}^2$	0.084	0.126	0.002	0.010	0.409	0.009	17.462	2.665	0.204
Wald stat	-	0.771	<b>442.1</b>	<b>167.2</b>	-	<b>924.5</b>	-	-	<b>13078</b>
$p$ -value	-	0.380	0	0	-	0	-	-	0

Bold values denote significance at the 10% level against the null that  $R^2$  from  $r_{t+q}(q) = \alpha_q + \beta_q r_t(q) + \varepsilon_{t+q}$  is less than or equal to its theoretical upper bound  $\bar{R}^2$ .  $p$ -values are asymptotic.

test at the 10% level could never reject that the variance ratio is different from one against a negatively serially correlated alternative from the 28 quarter holding period onwards. This pattern is consistent across the two sub-samples. The finding from the whole sample - that the variance ratios indicated negative serial correlation in the Cecchetti-Lam-Mark residuals but are not significant - is replicated.

As in the whole sample, there is a small number of rejections from the quantilegram test in both sub-samples. In general, the  $p$ -values are relatively large. There is thus little evidence against the Cecchetti-Lam-Mark model in the sub-sample quantilegrams. Nonetheless, the first sub-sample gives rise to a significant rescaled range statistic, although the second does not. Yet again, this shows the value of applying different test statistics. It seems the first sample residuals exhibit a type of long-range dependence not detected in the variance ratios or quantilegrams.

The maximal predictability results do appear to be sensitive to the sample used. While there are three significant violations in the first period and four in the second, these come at different horizons. Only for  $q = 28$  is the  $R^2$  bound violated in both sub-samples. Curiously, it is not violated in the whole sample for this holding period. In the whole sample, the only  $R^2$  bound violation is for  $q = 24$ . This exceedence is highly significant. However, the  $R^2$  bound is violated only in the first sub-sample at this horizon. Both sub-samples show some very large  $R^2$ s, especially for horizons of 32 quarters or more. This is another pattern not reflected across the whole sample.

The differences in the holding periods for which the  $R^2$  bound is violated between the two sub-samples can be interpreted in two ways. The first is to argue that the instability of the exceedence occurrences over time suggests that the results are due, at least in part, to sampling error. In this interpretation, the evidence in Table 13 strengthens that in Table 10 that the high/low growth state variable provides a reasonable explanation of the degree of predictability observed, in that the degree of predictability observed is not “too much”. However, it appears that the exceedences of the  $R^2$  bounds in the sub-samples are too frequent and too significant to be explained by sampling error alone. But, Table 13 suggests that the high/low growth indicator cannot explain the serial



dependence structure of returns alone. That the indicator can explain the degree of predictability observed in some periods, but not others, and at some horizons and not others is consistent with evidence which suggests that the factors that correlate with the risk premium change over time (e.g. Li and Tsiakis, 2016).

## 6.2 Robustness

In this section we evaluate the robustness of our main results to (i) concerns about look-ahead bias, (ii) using the identity weight matrix in estimation of the asset pricing models and (iii) using an annual time horizon. Overall, the conclusions drawn above are robust to look-ahead bias considerations and the choice of weight matrix and decision interval. In terms of look-ahead bias, the Campbell-Cochrane evidence is unchanged. The form of the departure of the Bansal-Yaron residuals changes once look-ahead bias is accounted for, but the conclusion that they are not MDS remains. The evidence regarding the Cecchetti-Lam-Mark model is similar, but stronger against it. For the Bansal-Yaron and Campbell-Cochrane models, the form of the violation of the martingale difference null seems to change depending on the weight matrix or data frequency, but the evidence of violation remains. For the Cecchetti-Lam-Mark model, the evidence of violation and form of violation remains similar, although the rescaled range statistic becomes significant when we use annual data.

### 6.2.1 Look-ahead bias

A concern about the above results may be that the parameters of the ex-ante ( $t - j$ ) expectations are estimated over future ( $t - j + 1, t - j + 2, \dots$ ) data, which could induce a bias in the test statistics. Note that these concerns apply on to the variance ratio and rescaled range tests. The quantile bootstrap procedure explicitly accounts for the estimation method and the maximum predictability test conditions on the parameter estimates. We evaluate the robustness of our variance ratio and rescaled range results to using past data only to estimate the parameters of the models. We compute residuals for the second sub-sample only which are formed using parameters estimated over an expanding sample beginning at the first observation and ending at the  $(t - j)$ th observation. We compare these results to those obtained for the second sub-sample above to evaluate the effect of restricting the data sample to past data only instead of past and future data. We have already seen that the conclusions of the test statistics are robust to which sub-sample is used and whether we use the full sample or a sub-sample.

The Bansal-Yaron model rescaled range statistic is very similar to that in the second period. The variance ratios are rather different, however. While all the variance ratios for  $q = 12$  onwards are significant, they are less than, not greater than, one. Nonetheless, the conclusion that the residuals are not MDS remains intact. For the Campbell-Cochrane model, the variance ratio and rescaled range results are very similar to the second sub-sample. The Cecchetti-Lam-Mark variance ratio results are also very similar, although the rescaled range test is just about significant at 1.75 ( $p$ -value 0.10), so that the evidence against the Cecchetti-Lam-Mark model becomes stronger.

### 6.2.2 Identity weight matrix

The Bansal-Yaron quantile bootstrap, rescaled range and maximal predictability results are robust to using the identity weight matrix, albeit with slightly fewer significant violations

in the latter. The significance of the quantilograms is also almost identical to the optimal weight matrix case, too. However, the pattern of most of the significant quantilograms being greater than the mean of the bootstrapped distribution is reversed. The conclusion that the Bansal-Yaron residuals are not MDS remains. The problem with many of the  $R^2$  bounds being either greater than one or less than zero remains. The variance ratio results are rather different to the optimal weight matrix case. They are all less than one and not significant. Nonetheless, the quantilograms and rescaled range statistics provide robustly significant evidence that the Bansal-Yaron model cannot explain the serial dependence structure of the market return. The model remains obviously mis-specified in terms of the mean residual.

The Campbell-Cochrane rescaled range result is robust to using the identity matrix in estimation. The quantilogram results provide a stronger rejection of the null that the model residuals are MDS, as there are many more rejections (63). These all occur for  $q \geq 16$ . The maximal predictability test shows a significant  $\bar{R}^2$  exceedence for  $q = 24$  only, despite there being exceedences at  $q = 16$  and  $q = 20$ . However, as when using the optimal weight matrix, there are significant exceedences of  $\bar{R}^2$  for  $q = 3, 4, 5, 6$  quarters, reinforcing the point that the problems the Campbell-Cochrane model has in explaining the amount of predictability observed are concentrated at shorter horizons. The variance ratios are similar to the optimal weight matrix case, with a mix of values both less than and greater than one. There are only two rejections (at  $q = 16$  and  $q = 36$ ). Thus the rejection of the Campbell-Cochrane model's ability to explain the serial dependence properties of the market return is robust, although the violation takes a different form using the identity matrix. It is clear that the Campbell-Cochrane model cannot explain the serial dependence structure of the market return using quarterly data.

The Cechetti-Lam-Mark variance ratio evidence is almost identical when the identity weight matrix is used. There is again evidence of negative autocorrelation in the residuals. The quantilogram evidence is also very similar to the optimal weight matrix case. There are 10 significant violations in total. The rescaled range statistic has a value of 1.5, similar to when the optimal weight matrix is used. The maximal predictability evidence is the same whether the identity or optimal weight matrix is used, since the state variable is not estimated by GMM.

### 6.2.3 Annual decision interval

The Bansal-Yaron results are very much robust to using an annual decision interval. The only difference is that the rescaled range statistic rejection now comes in the upper tail of the distribution ( $\hat{Q}^K = 1.93$ ,  $p$ -value: 0.03), so that the residuals show persistent long-range dependence. All of the variance ratios are significantly different from one and greater than one. Most significance is at the 1% level. The quantilogram results are very similar to those at the quarterly frequency. In fact, they provide an even stronger rejection of the Bansal-Yaron model's ability to explain the serial dependence properties of the market return, since there are also several rejections of the null at at the 99th percentile. Again, the significant weighted quantilograms are mainly greater than the mean of the bootstrap distribution. The modified rescaled range statistic has a value of 18, and provides a further overwhelming rejection of the null. There is a similar number of significant violations of the  $R^2$  bound in the maximal predictability test, but they occur at different horizons (in terms of years) to in the quarterly data. The problem with many  $R^2$  bounds being either greater than one or less than zero remains.

The Cecchetti-Lam-Mark model results are also similar when moving to an annual decision interval. The change is that the rescaled range is now significant ( $\widehat{Q}^K = 1.96$ ,  $p$ -value: 0.03), so that the martingale difference null is outright rejected. The variance ratios are all less than one, have a “U” shape in terms of the point estimates and are not significantly different from one. Again, the standard errors are fairly large, and the evidence suggests some negative serial correlation in the residuals. The quantilograms also tell a similar story as at the quarterly frequency, except there are now more rejections (11) and, again, it appears that there is some non-linear predictability in the residuals. There is still only one significant  $R^2$  bound violation in the maximal predictability test. This is at a different holding period to in the quarterly results.

The annual variance ratios and quantilogram results are somewhat different for the Campbell-Cochrane model. The variance ratios are less than one. None are significant at the 10% level. However some are quite different from one, the smallest being 0.54, and the standard errors are relatively large. There is thus some evidence of negative serial correlation and the low power of the variance ratio statistic against such alternatives. The quantilogram results only show significant violations of the null that the model residuals are MDS at the 1st, 5th, 95th and 99th percentiles. I.e. rejections are concentrated in extreme quantiles. This again highlights the value of using multiple test statistics. While these results still provide clear evidence against the Campbell-Cochrane model, it is not as strong as in the quarterly data. The rescaled range statistic is insignificant.

## 7 Economic importance of predictability

Having shown above that (a) there appears to be some evidence of mean reversion in the market and (b) that the models of asset prices we consider cannot explain this, we turn to the economic significance of the predictability observed and the failure of the models to explain it. There is strong evidence that a timing strategy exploiting information in the serial correlations of market returns can produce significant economic gains to investors under the Bansal-Yaron model. There is also some evidence that, under the Campbell-Cochrane model, the economic losses from the timing strategy are not as large as they ought to be. While the economic gains found under the Cecchetti-Lam-Mark model are consistent with the model, the certainty equivalent returns produced by the model are unusually large in magnitude, due to the extreme estimated risk aversion.

If there is long-run serial correlation in the market, then, for a suitably large  $q$ , lagged returns  $r_{m,t-1}, r_{m,t-2}, \dots, r_{m,t-q}$  are able to forecast  $r_{m,t}$ . One way to construct such a forecast is from the regression

$$r_{m,t} = a + b_1 r_{m,t-1} + b_2 r_{m,t-2} + \dots + b_q r_{m,t-q} + e_t. \quad (24)$$

We denote by  $\hat{r}_{m,t}$  the forecast of  $r_{m,t}$  given  $\{r_{m,s}\}_{s=1}^{t-1}$ , i.e.

$$\hat{r}_{m,t} = \hat{a} + \sum_{j=1}^q \hat{b}_j r_{m,t-j},$$

where  $\hat{a}, \hat{b}_1, \dots, \hat{b}_q$  are the OLS estimates from (24).

A long-only market timing strategy to exploit this predictability is as follows:

- buy into the market at  $t$  if  $\hat{r}_{m,t+1}$  is greater than the median of  $\{r_{m,s}\}_{s=1}^t$ ,

- sell out of the market and buy bonds with the proceeds if  $\hat{r}_{m,t+1}$  is less than the first decile of  $\{r_{m,s}\}_{s=1}^t$ ,
- re-invest any dividends but otherwise maintain the previous period's position if neither of the above conditions are met.

The long-short version of this strategy would short the market instead of buying bonds when  $\hat{r}_{m,t+1}$  is less than the first decile of  $\{r_{m,s}\}_{s=1}^t$ . In most periods, the market outperforms bonds and going short is somewhat riskier than taking a long position, as the losses to the former are potentially unlimited. Hence, our strategy has a strong preference for being in the market. Evaluating the strategy over the same range of  $q$  as the variance ratios, our strategy is typically out of the market about 4-10% of the time, depending on  $q$ .

In practice, buying the market index is easily achieved by buying an ETF which tracks the market portfolio. These ETFs can be shorted, or inverse ETFs can be bought when the investor wishes to short the market. The bonds considered are US one-month Treasury bills. As the data frequency is quarterly, we assume that the investor rolls over three months of one-month T-bills.

We consider strategies of the form above and compare certainty equivalent return of an investor who follows the strategy outline to that of one who simply holds the market portfolio. Note that this is not necessarily an optimal strategy and does not allow for optimal diversification between the market timing strategy and the market index. The certainty-equivalent gain from using the timing strategy is thus understated.

We calculate certainty equivalents on a two-period basis, i.e. for an agent who lives for two quarters  $t$  and  $t+1$ . We also assume that the strategy return at  $t+1$  given information available at  $t$ ,  $R_{t+1}|\mathcal{F}_t$  is lognormally distributed. This permits analytical solutions to the certainty equivalents. In addition, the agent consumes only at  $t+1$  and invests all of her wealth in either the timing strategy or the buy-and-hold strategy. The certainty equivalent wealth will be  $W_{t+1}^{CE}$  such that  $U(C_t = 0, C_{t+1} = W_{t+1}) = E_t(U(C_t = 0, W_{t+1}))$ , where  $W_{t+1} = (1 + R_t)W_t$ . The certainty equivalent is calculated as if the investor's only source of consumption  $t+1$  is the wealth invested at  $t$ .  $W_0$  is normalised to be 1. The certainty equivalent return is  $(1 + R_{t+1}^{CE}) = W_{t+1}^{CE}/W_t$ . We consider the sample mean log certainty equivalent return,  $E(r_{t+1}^{CE}) \equiv E(\Delta w_{t+1}^{CE})$ . Of course, this is approximately equal to  $E(R_t^{CE})$ .

This approach is a little rough, and so is only illustrative of the economic importance of the predictability in the market return. However, this period-by-period approach and the use of distributional assumptions regarding returns is fairly standard in the literature and has been used, *inter alia*, by DeMiguel et al. (2007). Where we allow for transactions costs, we also follow the standard procedure (used by, e.g. Brandt et al., 2009; DeMiguel et al., 2007) in assuming a proportional transaction cost of 50bp each way. Note that we do not impose any transaction costs on the buy-and-hold portfolio: the 50bp each-way cost is the additional cost of the active strategy.

The certainty equivalents are given as follows. For the Campbell-Cochrane model, we have

$$\Delta w_{t+1}^{CE} = E_t r_{t+1} + \frac{1}{2}(1 - \phi)^2 [\text{Var}_t r_{t+1} + \lambda^2 (s_t) \sigma_\nu^2 + 2 \text{Cov}_t(r_{t+1}, s_{t+1})] + \frac{1}{2}(1 - \gamma) \lambda^2 (s_t) \sigma_\nu^2.$$

It is then straightforward to compute  $E(\Delta w_{t+1}^{CE})$ . Assuming  $R_{t+1}|\mathcal{F}_t$  to be lognormally

distributed,

$$\begin{pmatrix} w_{t+1} \\ s_{t+1} \end{pmatrix} \Big| \mathcal{F}_t \sim N \left( \begin{pmatrix} w_t + r_{t+1} \\ (1 - \phi)\bar{s} + \phi s_t \end{pmatrix}, \begin{pmatrix} \text{Var}_t r_{t+1} & \text{Cov}_t(r_{t+1}, s_{t+1}) \\ \text{Cov}_t(r_{t+1}, s_{t+1}) & \lambda^2(s_t)\sigma_\nu^2 \end{pmatrix} \right).$$

Given  $U(C_t = 0, C_{t+1} = W_{t+1}) \propto (S_{t+1}W_{t+1})^{1-\gamma}$  and the properties of the multivariate lognormal distribution, we arrive at the above result.

For the Cecchetti-Lam-Mark and Bansal-Yaron models, we have

$$\Delta w_{t+1}^{CE} = E_t r_{t+1} + \frac{1}{2}(1 - \gamma) \text{Var}_t r_{t+1}. \quad (25)$$

For the Cecchetti-Lam-Mark model, this follows directly from  $U(C_t = 0, C_{t+1} = W_{t+1}) \propto W_{t+1}^{1-\gamma}$  and the properties of the lognormal distribution. For the Bansal-Yaron model, the utility index when  $C_t = 0$  and  $C_{t+1} = W_{t+1}^{CE}$  with certainty is given by

$$V_t = \delta \left( (1 - \delta)^{\frac{\theta}{1-\gamma}} W_{t+1}^{CE} \right).$$

Since there is no period  $t + 2$ ,  $V_{t+i} = 0$  for  $i > 2$ . For a stochastic  $W_{t+1}$ ,

$$V_t = \delta \left( (1 - \delta)^{\frac{\theta}{1-\gamma}} (E_t(W_{t+1}^{1-\gamma}))^{\frac{1}{1-\gamma}} \right).$$

Therefore,  $W_{t+1}^{CE} = (E_t(W_{t+1}^{1-\gamma}))^{\frac{1}{1-\gamma}}$  and combining this with the lognormality of  $W_{t+1}$  yields (25).

The above largely ignores the consumption and dividend endowment processes assumed by the models considered, except insofar as they affect the estimation of the relevant preference parameters. The reason for this is that these models are closed as endowment representative-agent economies. It is therefore not, strictly speaking, possible for the investor to earn anything other than the aggregate market return. This is not a situation observed in practice. Since this test relates to the practical economic significance of the findings in this paper, we focus on the preferences implied by these models.

Our null hypothesis is that there are no certainty equivalent gains from the timing strategy. If the model in question accounts for the entire serial correlation structure of the market, a strategy which uses information from the serial correlations of the market return should not be able to outperform the market on a risk-adjusted basis. To test this null, we use the following bootstrap procedure.

1. Resample the residuals  $\xi_t = r_{m,t} - E_{t-1} r_{m,t}$ , where  $E_{t-1} r_{m,t}$  is the model-implied expected return, by applying wild bootstrap, using the Mammen two-point distribution, to them. Call these replicate residuals  $\xi_t^*$ .
2. Construct  $r_{m,t}^* = E_{t-1} r_{m,t} + \xi_t^*$ . Conditioning on the dividend and consumption data, use these  $r_{m,t}^*$  to construct the commensurate log price-dividend ratio series,  $z_{m,t}^*$ .
3. Re-estimate the model based on  $r_{m,t}^*$ ,  $z_{m,t}^*$  and the dividend and consumption data.
4. Generate ex-post market returns from the model.<sup>21</sup>

---

<sup>21</sup>For the Bansal-Yaron model, the Campbell-Shiller log-linearisation on which Constantinides and Ghosh (2011) base their inversion of the model gives the equilibrium log market return as a function of observables and  $\kappa_{0,m}$  and  $\kappa_{1,m}$ . The Campbell-Cochrane ex-post returns can be calculated from (12). Cecchetti et al. (1990) provide a formula for the equilibrium market return in terms of the high/low growth state variable and consumption.

Table 14: Timing strategy performance

$q$	4	8	12	16	20	24	28	32	36	40
Market										
Mean (%)	1.83	1.76	1.55	1.54	1.50	1.46	1.49	1.55	1.77	1.92
Std dev (%)	8.21	8.34	8.42	8.49	8.62	8.63	8.74	8.64	8.72	8.19
Sharpe ratio	0.191	0.175	0.148	0.143	0.136	0.133	0.135	0.145	0.165	0.189
One-month T-bill										
Mean (%)	0.26	0.30	0.31	0.32	0.32	0.31	0.31	0.30	0.33	0.38
Std dev (%)	0.62	0.58	0.59	0.59	0.61	0.62	0.64	0.66	0.66	0.66
Long-only, no transactions costs										
Mean (%)	1.69	1.63	1.36	1.32	1.38	1.47	1.41	1.60	1.80	2.13
Std dev (%)	8.14	8.28	8.26	8.28	8.47	8.33	8.51	8.53	8.58	7.66
Sharpe ratio	0.175	0.161	0.127	0.121	0.125	0.139	0.129	0.151	0.171	0.229
Long-only, with transactions costs										
Mean (%)	1.69	1.59	1.30	1.23	1.24	1.30	1.23	1.44	1.61	1.93
Std dev (%)	8.14	8.29	8.27	8.29	8.48	8.34	8.53	8.63	8.59	7.66
Sharpe ratio	0.175	0.156	0.120	0.109	0.108	0.118	0.108	0.133	0.149	0.203
Long-short, no transactions costs										
Mean (%)	1.76	1.49	1.14	1.04	1.25	1.47	1.32	1.62	1.76	2.22
Std dev (%)	8.23	8.39	8.49	8.57	8.66	8.63	8.77	8.62	8.73	8.12
Sharpe ratio	0.182	0.142	0.098	0.084	0.107	0.134	0.116	0.153	0.164	0.227
Long-short, with transactions costs										
Mean (%)	1.76	1.47	1.11	1.00	1.22	1.43	1.29	1.60	1.72	2.22
Std dev (%)	8.23	8.40	8.50	8.59	8.67	8.64	8.78	8.63	8.74	8.13
Sharpe ratio	0.182	0.140	0.095	0.079	0.104	0.129	0.112	0.150	0.159	0.220

All returns are continuously compounded real returns at the quarterly frequency. In particular, the one-month T-bill return is from three rolled over one-month T-bills.

5. Re-run the timing and buy-and-hold strategies on the model-implied ex-post returns.
6. Use the returns and preference parameters from step 3 to calculate the certainty-equivalent return gain from the timing strategy.

This bootstrap automatically accounts for any finite-sample bias induced by the estimation of the asset pricing models.

Table 14 shows that the timing strategy always out-performs the market for  $q = 40$  in terms of mean returns and Sharpe ratios. There is also some other outperformance of the market, even when transactions costs are accounted for, for  $q = 24$  and  $q = 32$ . The long-only version of the strategy also has a lower standard deviation than the market. In the context of the models above, these metrics are not greatly meaningful. The object of interest to an investor is the certainty-equivalent return, which automatically adjusts the returns for risk according to the investor's utility function. The standard deviation and Sharpe ratio account for only specific types of risk.

Table 15: Certainty equivalent of the timing strategy

$q$	4	8	12	16	20	24	28	32	36	40
Long-only, no transactions costs										
BY-tim	1.45	1.34	1.06	1.03	1.08	1.17	1.10	1.29	1.48	1.87
BY-b/h	1.56	1.45	1.28	1.24	1.16	1.16	1.12	1.29	1.42	1.58
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CC-tim	1.73	1.63	1.35	1.31	1.38	1.47	1.40	1.59	1.79	2.12
CC-b/h	1.84	1.74	1.58	1.54	1.48	1.47	1.45	1.60	1.74	1.86
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CLM-tim	-81.2	-83.0	-82.8	-83.1	-87.0	-84.0	-87.7	-88.0	-88.8	-70.1
CLM-b/h	-81.3	-84.0	-85.9	-87.3	-90.1	-90.3	-93.1	-90.7	-92.2	-81.4
$p$ -value diff	0.23	0.23	0.23	0.23	0.23	0.23	0.27	0.27	0.34	0.34
Long-only, with transactions costs										
BY-tim	1.40	1.28	0.99	0.92	0.91	0.99	0.90	1.11	1.28	1.66
BY-b/h	1.56	1.45	1.28	1.24	1.16	1.16	1.12	1.29	1.42	1.58
$p$ -value	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CC-tim	1.70	1.58	1.29	1.22	1.23	1.30	1.22	1.43	1.60	1.92
CC-b/h	1.84	1.74	1.58	1.54	1.48	1.47	1.45	1.60	1.74	1.86
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CLM-tim	-77.2	-80.1	-80.1	-80.6	-84.3	-81.5	-85.2	-85.1	-86.0	-67.9
CLM-b/h	-81.3	-84.0	-85.9	-87.3	-90.1	-90.3	-93.1	-90.7	-92.2	-81.4
$p$ -value	0.28	0.28	0.28	0.28	0.28	0.28	0.34	0.34	0.38	0.38
Long-short, no transactions costs										
BY-tim	1.47	1.19	0.83	0.73	0.93	1.15	1.00	1.31	1.43	1.93
BY-b/h	1.56	1.45	1.28	1.24	1.16	1.16	1.12	1.29	1.42	1.58
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CC-tim	1.75	1.49	1.13	1.04	1.24	1.46	1.32	1.62	1.75	2.20
CC-b/h	1.84	1.74	1.58	1.54	1.48	1.47	1.45	1.60	1.74	1.86
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CLM-tim	-80.9	-84.5	-86.8	-88.6	-90.2	-89.4	-92.4	-89.0	-91.1	-78.1
CLM-b/h	-81.3	-84.0	-85.9	-87.3	-90.1	-90.3	-93.1	-90.7	-92.2	-81.4
$p$ -value diff	0.34	0.34	0.34	0.34	0.34	0.34	0.37	0.37	0.42	0.42
Long-short, with transactions costs										
BY-tim	1.27	1.13	0.75	0.62	0.76	0.96	0.80	1.13	1.23	1.71
BY-b/h	1.56	1.45	1.28	1.24	1.16	1.16	1.12	1.29	1.42	1.58
$p$ -value	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CC-tim	1.57	1.45	1.08	0.95	1.10	1.29	1.14	1.46	1.56	2.00
CC-b/h	1.84	1.74	1.58	1.54	1.48	1.47	1.45	1.60	1.74	1.86
$p$ -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CLM-tim	-79.8	-82.6	-85.0	-87.0	-88.3	-87.6	-90.7	-86.9	-89.1	-76.3
CLM-b/h	-81.3	-84.0	-85.9	-87.3	-90.1	-90.3	-93.1	-90.7	-92.2	-81.4
$p$ -value	0.22	0.23	0.23	0.23	0.23	0.23	0.27	0.27	0.30	0.30

Certainty equivalents given as percentages.  $p$ -values are bootstrapped tests of the null that the certainty equivalent return of the timing strategy is greater than or equal to that of the buy and hold strategy. BY-tim denotes the certainty equivalent return of the timing strategy under the Bansal-Yaron model, BY-b/h does likewise for the buy and hold strategy. CC and CLM denote the Campbell-Cochrane and Cecchetti-Lam-Mark models, respectively.

Table 15 shows that, for the Bansal-Yaron model, the timing strategy produces some estimated certainty equivalent return gains for higher  $q$ . Only the gains for  $q = 40$  survive once transaction costs are accounted for. These gains are relatively small, amounting to about 0.3% per year. Crucially, though, almost all of the Bansal-Yaron certainty equivalent gains are significantly too large, even when they are negative. Even where the timing strategy leads to a lower certainty equivalent return for the investor than buy and hold, this lower return is not sufficiently lower than the buy-and-hold return to be consistent with the Bansal-Yaron model. This finding does survive the inclusion of transaction costs.

For the Campbell-Cochrane model, while certainty equivalent gains are produced for  $q = 40$  for all implementations of the strategy (long-only/long-short, with/without transactions costs), these are not significant. The only significant certainty equivalent gain for a timing strategy under the Campbell-Cochrane model is for  $q = 4$  in the long-short strategy with transaction costs. This “gain” is actually negative, so, again, the timing strategy does not underperform the market as much as it should on a risk-adjusted basis.

Some of the certainty equivalent gains, produced under the Cecchetti-Lam-Mark model are very large. The difference in certainty equivalent returns between the timing and buy-and-hold strategies can be interpreted as the largest additional fee an investor would pay to have access to the timing strategy. Our results suggest that an investor would be willing to pay fees of up to 44% a year to access the timing strategy. Yet none of these certainty equivalent gains is significant. However, the Cecchetti-Lam-Mark results in Table 15 are rather too large in magnitude to be a plausible financial time series, principally due to the extreme estimated risk aversion. We do not attach too much weight to them.

## 8 Conclusion

In this paper, we use GMM to estimate three consumption-based models of asset prices, the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models, and derive ex-ante expected market returns from them. If the model in question is well specified, and can account for the serial correlation structure of the market return, the difference between the market return and the model-implied expected return is MDS. To test this hypothesis, we modify the variance-ratio and rescaled range statistics to account for these expected returns being functions of estimated parameters. We also use a weighted quantilogram test based on a bootstrap procedure robust to the estimation of model parameters.

We find that the difference between the market return and the model-implied expected returns is not MDS and thus that the models are not well specified and cannot account for the dynamic structure of the market returns. This finding is robust to the sample period, GMM weight matrix and decision interval used. The Bansal-Yaron and Campbell-Cochrane residuals produce highly significant rescaled range statistics and there is evidence that they are serially correlated in levels and in quantile hits. The main significant evidence against the null that the Cecchetti-Lam-Mark model residuals are MDS comes from the rescaled range statistics in the first sub-sample, the second sub-sample once look-ahead bias is accounted for and at the annual frequency. However, its residuals produce variance ratios that are rather below one but with large standard errors for



all the weight matrices and decision intervals considered. Coupled with the known low power of variance ratio (and other tests of the martingale difference hypothesis) against a negatively serially correlated alternative (Poterba and Summers, 1988), the variance ratios strongly suggest that the Cecchetti-Lam-Mark residuals are not MDS.

We also present evidence that it is possible to outperform the market on a risk-adjusted basis under the Bansal-Yaron model, using information from the serial correlations of returns. This would not be possible if the Bansal-Yaron accounted for those serial correlations. While outperformance of the market using a timing strategy is also possible under the Campbell-Cochrane model, this outperformance is not significant. In one case, however, the risk-adjusted underperformance of the timing strategy is significantly less than it ought to be under the Campbell-Cochrane model. It is possible to outperform the market substantially in certainty equivalent terms under the Cecchetti-Lam-Mark model, but this outperformance is not significant. We attach little weight to these results since the Cecchetti-Lam-Mark certainty equivalent returns do not form an economically plausible financial time series.

For the Bansal-Yaron and Campbell-Cochrane models, an  $R^2$  test of maximal predictability produces evidence that the amount of time-series predictability in the market return is more than can be explained by the state variables used in the stochastic discount factor. This evidence is stronger in the sub-samples. We do, however, note some empirical difficulties in implementing this test.

The failures of the models tested to adequately capture the serial correlation structure of the market return suggest several possible avenues of research. One would be to modify the sets of state variables used by the models we consider, or to come up with new state variables altogether. Alternatively, given that the models we consider here are models of equilibrium stock prices and returns, it may be necessary to embed the models we consider in a framework that models adjustment to the equilibrium. This would generate natural mean reversion Adam et al. (2016) present such a model. They have an agent with CRRA preferences who knows the risk-adjusted price is a random walk (a result due to Samuelson, 1965) but who observes the risk-adjusted price plus mean-zero noise. Optimal updating of beliefs under subjective expected utility maximisation produces a feedback loop: expectations affect prices, as in the classical model, but prices also affect expectations, due to updating. This feedback imparts upon the prices mean reversion and excess volatility, even though the estimated prior uncertainty (noise variance) is small. In general, this model is able to match many facts about asset prices, including the long-horizon predictability of excess returns with respect to the price-dividend ratio. However, rather like the standard CRRA model, it cannot capture the equity premium. Nonetheless, it is possible that by applying this framework to, say, the Campbell-Cochrane model we could also capture the equity premium.

A different model of adjustment to equilibrium accounts for transaction costs. For example, Seo (2003) takes the expectations hypothesis as the equilibrium model of the bond market. However, arbitrage opportunities are only taken advantage of if the profit they yield is greater than the cost of the transaction. Encoding this modified expectations hypothesis in a threshold vector error correction model, Seo finds significant threshold effects and rejects the vanilla expectations hypothesis in favour of the augmented expectations hypothesis with transaction costs. He also rejects pure preferred habitat theory against an expectations hypothesis alternative. As the models we consider above imply equilibrium stock prices, it may be possible to explain the dynamic properties of the market return using a variant of this threshold methodology. Of course, it may also be

that an “outright” behavioural model, which goes beyond rational learning, is required.

## Appendices

### A Bansal-Yaron model estimation

#### A.1 Inversion and stochastic discount factor coefficients

Constantinides and Ghosh (2011) show that

$$\begin{aligned}x_t &= \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 z_{m,t} \\ \sigma_t^2 &= \beta_0 + \beta_1 r_{f,t} + \beta_2 z_{m,t},\end{aligned}$$

where

$$\begin{aligned}\alpha_0 &= \frac{A_{2,m}A_{0,f} - A_{0,m}A_{2,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\ \alpha_1 &= \frac{-A_{2,m}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\ \alpha_2 &= \frac{A_{2,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\ \beta_0 &= \frac{A_{0,m}A_{1,f} - A_{1,m}A_{0,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\ \beta_1 &= \frac{A_{1,m}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\ \beta_2 &= \frac{-A_{1,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}}.\end{aligned}$$

The expressions for the  $A_0, \dots, A_{2,f}$  coefficients are given by

$$\begin{aligned}
A_1 &= \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho_x} \\
A_2 &= \frac{\frac{1}{2} \left[ \left( -\frac{\theta}{\psi} + \theta \right)^2 + (\theta \kappa_1 A_1 \psi_x)^2 \right]}{\theta(1 - \kappa_1 \nu)} \\
A_0 &= \frac{\ln(\delta) + \left( 1 - \frac{1}{\psi} \right) \mu_c + \kappa_0 + \kappa_1 A_2 \sigma^2 (1 - \nu) + \frac{1}{2} \theta \kappa_1^2 A_2^2 \sigma_w^2}{1 - \kappa_1} \\
A_{1,m} &= \frac{\phi - \frac{1}{\psi}}{1 - \kappa_{1,m} \rho_x} \\
A_{2,m} &= \frac{(1 - \theta) A_2 (1 - \kappa_1 \nu) + \frac{1}{2} [\gamma^2 + \varphi^2 + ((\theta - 1) \kappa_1 A_1 + \kappa_{1,m} A_{1,m})^2 \psi_x^2]}{1 - \kappa_{1,m} \nu} \\
A_{0,m} &= \frac{\theta \ln(\delta) + \left( -\frac{\theta}{\psi} + \theta - 1 \right) \mu_c + (\theta - 1) \kappa_0 + (\theta - 1) (\kappa_1 - 1) A_0 + (\theta - 1) \kappa_1 A_2 \sigma^2 (1 - \nu)}{1 - \kappa_{1,m}} \\
&\quad + \frac{\kappa_{0,m} + \mu_d + \kappa_{1,m} A_{2,m} \sigma^2 (1 - \nu) + \frac{1}{2} [(\theta - 1) \kappa_1 A_2 + \kappa_{1,m} A_{2,m}]^2 \sigma_w^2}{1 - \kappa_{1,m}} \\
A_{0,f} &= \theta \ln(\delta) - \left( -\frac{\theta}{\psi} + \theta - 1 \right) \mu_c - (\theta - 1) \kappa_0 - (\theta - 1) (\kappa_1 - 1) A_0 - (\theta - 1) \kappa_1 A_2 (1 - \nu) \sigma^2 \\
&\quad - 0.5 (\theta - 1)^2 \kappa_1^2 A_2^2 \sigma_w^2 \\
A_{1,f} &= - \left[ \left( \frac{\theta}{\psi} + \theta - 1 \right) + (\theta - 1) (\kappa_1 \rho_x - 1) A_1 \right] \\
A_{2,f} &= - \left[ (\theta - 1) (\kappa_1 \nu - 1) A_2 + \frac{1}{2} \left( \left( -\frac{\theta}{\psi} + \theta - 1 \right)^2 + (\theta - 1)^2 \kappa_1^2 + A_1^2 \psi_x^2 \right) \right].
\end{aligned}$$

In the stochastic discount factor

$$\exp \left\{ a_1 + a_2 \Delta c_{t+1} + a_3 \left( r_{f,t+1} - \frac{1}{\kappa_1} r_{f,t} \right) + a_4 \left( z_{m,t+1} - \frac{1}{\kappa_1} z_{m,t} \right) \right\} = 0,$$

we have:

$$\begin{aligned}
a_1 &= \theta \ln(\delta) + (\theta - 1) [\kappa_0 + (\kappa_1 - 1) (A_0 + A_1 \alpha_0 + A_2 \beta_0)] \\
a_2 &= -\frac{\theta}{\psi} + (\theta - 1) \\
a_3 &= (\theta - 1) \kappa_1 [A_1 \alpha_1 + A_2 \beta_1] \\
a_4 &= (\theta - 1) \kappa_1 [A_1 \alpha_2 + A_2 \beta_2].
\end{aligned}$$

The linearisation constants  $\kappa_0$  and  $\kappa_1$  derive from applying the Campbell and Shiller (1988) log-linearisation procedure to the returns to the consumption claim and market portfolio. These constants satisfy

$$\begin{aligned}
\kappa_1 &= \frac{\exp\{\bar{z}\}}{1 + \exp\{\bar{z}\}} \\
\kappa_0 &= \ln(1 + \exp\{\bar{z}\}) - \kappa_1 \bar{z},
\end{aligned}$$

where  $z_t$  is the log price/consumption ratio of an asset whose dividend stream is identical to consumption. Similar expressions are obtained for  $\kappa_{0,m}$  and  $\kappa_{1,m}$  when  $z$  is replaced by  $z_m$ . These are identified under the assumption that  $\bar{z}$  and  $\bar{z}_m$  are equal to the unconditional expectation of  $z_t$  and  $z_{m,t}$  respectively.

## A.2 Time-series moment conditions

The nine time-series moment conditions derived by Constantinides and Ghosh (2011) are:

$$\begin{aligned}
E(\Delta c_t) &= \mu_c \\
\text{Cov}(\Delta c_t, \Delta c_{t+1}) &= \rho_x \frac{\varphi_x^2 \sigma^2}{1 - \rho_x^2} \\
E(\Delta d_t) &= \mu_d \\
\text{Var}(\Delta d_t) &= \phi^2 \frac{\psi_x^2 \sigma^2}{1 - \rho_x^2} + \sigma^2 \varphi_u^2 \\
\text{Cov}(\Delta d_t, \Delta d_{t+1}) &= \phi^2 \rho_x \frac{\varphi_x^2 \sigma^2}{1 - \rho_x^2} \\
\text{Cov}(\Delta c_t, \Delta d_t) &= \phi \frac{\varphi_x^2 \sigma^2}{1 - \rho_x^2} \\
\text{Var} [(\Delta c_t)^2] &= \frac{3\varphi_x^4 \sigma_w^2 (1 + \pi \rho_x^2)}{(1 - \rho_x^4)(1 - \pi^2)(1 - \pi \rho_x^2)} + \frac{1}{1 - \rho_x^4} \left[ 2\sigma^4 + \frac{4\rho_x^2 \varphi_x^4 \sigma^4}{1 - \rho_x^2} \right] + 2\sigma^4 \\
&\quad + \frac{3\sigma_w^2}{1 - \pi^2} + 4\mu_c^2 \frac{\varphi_x^2 \sigma^2}{1 - \rho_x^2} + \frac{6\varphi_x^2 \sigma_w^2 \pi}{(1 - \pi^2)(1 - \pi \rho_x^2)} + \frac{4\varphi_x^2 \sigma^4}{1 - \rho_x^2} + 4\mu_c^2 \sigma^2 \\
\text{Var} [(\Delta d_t)^2] &= \phi^4 \left[ \frac{3\varphi_x^4 \sigma_w^2 (1 + \pi \rho_x^2)}{(1 - \rho_x^4)(1 - \pi^2)(1 - \pi \rho_x^2)} + \frac{2\sigma^4}{1 - \rho_x^4} + \frac{4\rho_x^2 \varphi_x^4 \sigma^4}{(1 - \rho_x^4)(1 - \rho_x^2)} \right] \\
&\quad + \frac{3\sigma_w^2 \varphi_u^4}{1 - \pi^2} + 4\mu_c^2 \frac{\varphi_x^2 \sigma^2}{1 - \rho_x^2} \phi^2 + \frac{6\varphi_x^2 \sigma_w^2 \pi \phi^2 \varphi_u^2}{(1 - \pi^2)(1 - \pi \rho_x^2)} + \frac{4\varphi_x^2 \sigma^4}{1 - \rho_x^2} \phi^2 \varphi_u^2 \\
&\quad + 2\sigma^4 \varphi_u^4 + 4\mu_d^2 \varphi_u^2 \sigma^2.
\end{aligned}$$

## A.3 Expected return coefficients

The expected market return in the Bansal-Yaron model is

$$E_t r_{m,t+1} = B_0 + B_1 x_t + B_2 \sigma_t^2,$$

where

$$\begin{aligned}
B_0 &= \kappa_{0,m} + (\kappa_{1,m} - 1)A_{0,m} + \mu_d + \kappa_{1,m}A_{2,m}(1 - \nu)\sigma^2 - 3\kappa_{1,m} \\
B_1 &= A_{1,m}(\kappa_{1,m}\rho_x - 1) + \phi \\
B_2 &= A_{2,m}(\kappa_{1,m}\nu - 1).
\end{aligned}$$

## B Cecchetti-Lam-Mark $\kappa(y_t)$

$$\kappa(y_t) = \begin{cases} \tilde{\delta}(1 - \tilde{\delta}\tilde{\alpha}_1(p + q - 1))/\Delta & , y_t = 0 \\ \tilde{\delta}\tilde{\alpha}_1(1 - \tilde{\delta}(p + q - 1))/\Delta & , y_t = 1, \end{cases}$$

where

$$\begin{aligned}\tilde{\delta} &= \delta \exp\{\alpha_0(1 - \gamma) + (1 - \gamma)^2 \sigma_{y_t}^2 / 2\} \\ \tilde{\alpha}_1 &= \exp\{\alpha_1(1 - \gamma)\} \\ \Delta &= 1 - \tilde{\delta}(p\tilde{\alpha}_1 + q) + \tilde{\delta}^2 \tilde{\alpha}_1(p + q - 1).\end{aligned}$$

## References

- Adam, K., Marcet, A., and Nicolini, J. P. (2016). Stock market volatility and learning. *Journal of Finance*, 71(1):33–82.
- Baltas, N. and Karyampas, D. (in press). Forecasting the equity risk premium: the importance of regime-dependent evaluation. *Journal of Financial Markets*.
- Bansal, R. and Yaron, A. (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509.
- Barroso, P., Boons, M., and Karehnke, P. (2017). Time-varying predictability of consumption growth, macro-uncertainty, and risk premiums. Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2933449](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2933449).
- Brandt, M. W., Santa-Clara, P., and Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22(9):3411–3447.
- Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(2):149–192.
- Campbell, J. Y. (2014). Empirical asset pricing: Eugene Fama, Lars Peter Hansen, and Robert Shiller. *Scandinavian Journal of Economics*, 116(3):593–634.
- Campbell, J. Y. and Cochrane, J. H. (1998). Appendix for “By force of habit: a consumption-based explanation of aggregate stock market behaviour”. Harvard University, Cambridge, MA; University of Chicago, Chicago, IL.
- Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: a consumption-based explanation of aggregate stock market behaviour. *Journal of Political Economy*, 107(2):205–251.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
- Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–227.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.
- Cecchetti, S. G., Lam, P.-S., and Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review*, 80(3):398–418.

- Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, 96(5):893–920.
- Cochrane, J. H. (2008). Financial markets and the real economy. In Mehra, R., editor, *Handbook of the Equity Risk Premium*, pages 237–330. Elsevier Science, Amsterdam, Netherlands, 1st edition.
- Cochrane, J. H. and Saá-Requejo, J. (2000). Beykond arbitrage: good deal asset price bounds in incomplete markets. *Journal of Finance*, 108(1):79–119.
- Constantinides, G. M. (2006). Understanding the equity risk premium puzzle. University of Chicago, IL. Available at <http://faculty.chicagobooth.edu/george.constantinides/documents/Premium%20Essay%202006.pdf>.
- Constantinides, G. M., Donaldson, J. B., and Mehra, R. (2002). Juniors can't borrow: a new perspective on the equity premium puzzle. *Quarterly Journal of Economics*, 117(1):269–296.
- Constantinides, G. M. and Ghosh, A. (2011). Asset pricing tests with long-run risks in consumption growth. *Review of Asset Pricing Studies*, 1(1):96–136.
- DeMiguel, V., Garlappi, L., and Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953.
- Donaldson, J. and Mehra, R. (2008). Risk-based explanations of the equity premium. In Mehra, R., editor, *Handbook of the Equity Risk Premium*, pages 37–99. Elsevier Science, Amsterdam, Netherlands, 1st edition.
- Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behaviour of consumption and asset returns: a theoretical framework. *Econometrica*, 57(4):937–969.
- Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.
- Farmer, J. D. and Lo, A. W. (1999). Frontiers of finance: evolution and efficient markets. *Proceedings of the National Academy of Sciences*, 96(18):9991–2.
- Gangopadhyay, P. and Reinganum, M. R. (1996). Interpreting mean reversion in stock returns. *Quarterly Review of Economics and Finance*, 36(3):377–395.
- Garcia, R., Renault, E., and Semenov, A. (2004). A consumption CAPM with a reference level. Université de Montréal, Québec, Canada; available online at <http://epge.fgv.br/files/2056.pdf>.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

- Ho, C.-C. and Sears, R. S. (2006). Is there conditional mean reversion in stock returns? *Quarterly Journal of Business and Economics*, 45(3/4):91–112.
- Huang, D. and Zhou, G. (2015). Upper bounds on return predictability. Singapore Management University, Singapore; Washington University, St. Louis, MO. Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2426447](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2426447).
- Hurst, H. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799.
- Jegadeesh, N. (1991). Seasonality in stock price mean reversion: Evidence from the U.S. and the U.K. *Journal of Finance*, 46(4):1427–1444.
- Kandel, S. and Stambaugh, R. F. (1989). Modeling expected stock returns for long and short horizons. Department of Finance, Wharton School, University of Pennsylvania.
- Kim, C.-J., Morley, J. C., and Nelson, C. R. (2001). Does an intertemporal tradeoff between risk and return explain mean reversion in stock prices? *Journal of Empirical Finance*, 8(4):403–26.
- Kim, C.-J. and Nelson, C. R. (1998). Testing for mean reversion in heteroskedastic data II: Autoregression tests based on Gibbs-sampling-augmented randomization. *Journal of Empirical Finance*, 5(4):385 – 396.
- Kim, C.-J., Nelson, C. R., and Startz, R. (1998). Testing for mean reversion in heteroskedastic data based on Gibbs-sampling-augmented randomization. *Journal of Empirical Finance*, 5(2):131 – 154.
- Kim, M. J., Nelson, C. R., and Startz, R. (1991). Mean reversion in stock prices? A reappraisal of the empirical evidence. *Review of Economic Studies*, 58(3):515–528.
- Li, J. and Tsiakis, I. (2016). A new approach to equity premium prediction: Economic fundamentals matter! University of Guelph, Guelph, Canada.
- Linton, O. B. and Whang, Y.-J. (2007). The quantilogram: With an application to evaluating directional predictability. *Journal of Econometrics*, 141(1):250–282.
- Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica*, 59(5):1279–1313.
- Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66.
- Lo, A. W. and MacKinlay, A. C. (1989). The size and power of the variance ratio test in finite samples: a Monte Carlo investigation. *Journal of Econometrics*, 40(2):203–238.
- Malloy, C. J., Moskowitz, T. J., and Vissing-Jorgensen, A. (2009). Long-run stockholder consumption risk and asset returns. *Journal of Finance*, 64(6):2427–2479.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1):255–285.
- McQueen, G. (1992). Long-horizon mean-reverting stock prices revisited. *Journal of Financial and Quantitative Analysis*, 27(1):1–19.

- Mehra, R. and Prescott, E. C. (1985). The equity premium: a puzzle. *Journal of Monetary Economics*, 15(2):145 – 161.
- Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1):27–59.
- Richardson, M. and Smith, T. (1991). Tests of financial models in the presence of overlapping observations. *Review of Financial Studies*, 4(2):227–254.
- Richardson, M. and Stock, J. H. (1990). Drawing inferences from statistics based on multi-year asset returns. *NBER Working Paper*, 3335.
- RSAS (2013). Understanding asset prices. Compiled by the Economic Sciences Prize Committee of the Royal Swedish Academy of the Sciences (RSAS). Available at [http://www.nobelprize.org/nobel\\_prizes/economic-sciences/laureates/2013/advanced-economicsciences2013.pdf](http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2013/advanced-economicsciences2013.pdf).
- Salomons, R. (2008). A theoretical and practical perspective on the equity risk premium. *Journal of Economic Surveys*, 22(2):299–329.
- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49.
- Seo, B. (2003). Nonlinear mean reversion in the term structure of interest rates. *Journal of Economic Dynamics and Control*, 27.
- Siegel, J. J. (2005). Perspectives on the equity risk premium. *Financial Analysts Journal*, 61(6):61–73.
- Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(3):401 – 421.
- Wright, J. H. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economics Statistics*, 18(1):1–9.