UNIVERSITY OF CAMBRIDGE
Faculty of Economics

JANEWAY INSTITUTE CAMBRIDGE

CAMBRIDGE WORKING PAPERS IN ECONOMICS
JANEWAY INSTITUTE WORKING PAPERS

Do consumption-based asset pricing models explain the dynamics of stock market returns?

Michael Ashby
University of Cambridge

Oliver Linton
University of Cambridge

Abstract

We show that three prominent consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the dynamic properties of stock market returns. We show this by estimating these models with GMM, deriving ex-ante expected returns from them and then testing whether the difference between realised and expected returns is a martingale difference sequence, which it is not. Mincer-Zarnowitz regressions show that the models' out-of-sample expected returns are systematically biased. Furthermore, semi-parametric tests of whether the models' state variables are consistent with the degree of own-history predictability in stock returns suggest that only the Campbell-Cochrane habit variable may be able to explain return predictability, although the evidence on this is mixed.

Reference Details

2259	Cambridge Working Papers in Economics
2226	Janeway Institute Working Paper Series
Published	20 October 2022
Revised	03 August 2023
Key Words	consumption-based asset pricing models, serial correlation, predictability, martingale difference sequence, variance ratio, quantilogram, rescaled range, power spectrum, Mincer-Zarnowitz regression, MIDAS
JEL Codes	C52, C58, G12
Websites	www.econ.cam.ac.uk/cwpe www.janeway.econ.cam.ac.uk/working-papers

Do consumption-based asset pricing models explain the dynamics of stock market returns?*

Michael Ashby[†] Oliver Linton[‡]

August 2023

Abstract

We show that three prominent consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the dynamic properties of stock market returns. We show this by estimating these models with GMM, deriving ex-ante expected returns from them and then testing whether the difference between realised and expected returns is a martingale difference sequence, which it is not. Mincer-Zarnowitz regressions show that the models' out-of-sample expected returns are systematically biased. Furthermore, semi-parametric tests of whether the models' state variables are consistent with the degree of own-history predictability in stock returns suggest that only the Campbell-Cochrane habit variable may be able to explain return predictability, although the evidence on this is mixed.

JEL classification: C52, C58, G12
Keywords: consumption-based asset pricing models, serial correlation, predictability, martingale difference sequence, variance ratio, quantilogram, rescaled range, power spectrum, Mincer-Zarnowitz regression, MIDAS

*We thank Tiago Cavalcanti, Sonje Reiche, Melvyn Weeks and especially Mark Salmon, Donald Robertson and Gregory Connor for helpful comments and feedback, as well as participants at the Cambridge Econometrics Workshop. We are also very grateful to Yongmiao Hong and Yoon-Jin Lee for sharing the GAUSS code for implementing their generalised spectral test. Ashby acknowledges financial support from the Economic and Social Research Council under award 1515004.

[†]Downing College and Faculty of Economics, University of Cambridge. Regent Street, Cambridge, CB2 1DQ, UK. Email `mwa22@cam.ac.uk`.

[‡]Faculty of Economics and Trinity College, University of Cambridge. Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, UK. Email `obl20@cam.ac.uk`.

1 Introduction

Three prominent consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the dynamics of the US market return. The Bansal-Yaron and Campbell-Cochrane models are designed to explain the level of stock market returns, in particular to simultaneously resolve the equity premium and risk-free rate puzzles. Yet, whether these models can explain the degree of predictability in stock returns is of interest too, especially if investors want to time or beat the market. In this sense, the dynamics (second moment) of returns are important separately to their level (first moment). This is recognised by Cecchetti et al. (1990). The Cecchetti-Lam-Mark model was developed specifically to explain return dynamics, rather than to price assets per se.

Our first tests of whether the three models can explain return dynamics amount to testing whether the difference between the model-implied ex-ante expected market return and the realised market return - the residual - is a martingale difference sequence (MDS). Since the residuals are not MDS, there is some part of the dynamics of realised returns not captured by the models. To construct the expected returns and residuals, we first estimate the models by GMM. Our testing procedures account for this estimation step.

We base our tests of the null that the residuals are MDS on serial correlation, quantile hits, the rescaled range and the generalised spectrum of Hong (1999). The asymptotic distribution of the serial correlation and generalised spectrum-based tests accounts for the initial estimation step, while we use a bootstrap procedure to account for the estimation step in the quantile hits and rescaled range-based tests. We use a battery of tests since tests of the MDS null can suffer locally low power against certain alternatives (Poterba and Summers, 1988).

Our finding that none of the three models' residuals are MDS is robust to the empirical choices we make. It does not matter whether we use the optimal GMM weight matrix, or the identity matrix; whether we use size/book-to-market or industry portfolios to estimate the models; or whether we use quarterly, instead of annual, data. The only apparent hope comes from estimating the Cecchetti-Lam-Mark model using size/book-to-market portfolios and the identity GMM weight matrix at the quarterly frequency. However, using a quarterly sample gives a much larger number of observations and allows us to consider the robustness of our results over time by splitting the sample into two equal-length sub-samples. When we do this, we clearly reject the null that the Cecchetti-Lam-Mark residuals are MDS in both sub-samples.

As a further test, we use Mincer-Zarnowitz regressions (Mincer and Zarnowitz, 1969) to test whether out-of-sample model-generated expected returns are systematically biased, finding that they are. Rational expectations would rule out such systematic bias. We find significant evidence of systematic bias in the Bansal-Yaron expected returns. While the null of unbiased expected returns is not rejected for the Campbell-Cochrane and Cecchetti-Lam-Mark models in the main results which estimate the models using the optimal GMM weight matrix, size/book-to-market portfolios and annual data, there is a negative correlation between the expected returns and actual returns. In any case, the null of unbiasedness is rejected in all of the robustness check cases.

In each of the robustness check cases, we consider only models that provide credible expected returns. Many of the robustness check specifications do not give plausible expected returns series. This is less surprising than it might seem given the difficulties in identifying the parameters of asset pricing models (Cheng et al., 2022). There is no point checking the second moment of a model that fits poorly in terms of the first moment, as one would not use it to price assets anyway. Moreover, the centred second moment (e.g. serial correlation coefficient) is a function of the first moment.

We also consider semi-parametric tests of whether the degree of predictability in returns is consistent with the state variables of the three models being correctly specified. Unlike the residual-based tests, these tests do not depend on the functional form of the stochastic discount factor being correctly specified. They require only that the state variables be correctly specified.

Our first state-variable test is an adaptation of the Huang and Zhou (2017) test. We test whether the R^2 from a predictive regression of returns on their lagged values exceeds a theoretical upper bound, $\bar{R}^2$. $\bar{R}^2$ depends on the state variables of the stochastic discount factor (i.e. the state variables which explain stock returns).

Our second state-variable test comes from the Merton (1973) intertemporal CAPM (ICAPM). Merton shows that, if the ICAPM holds (for any risk-averse von Neumann-Morgenstern utility function), returns at $t+1$, r_{t+1}, can be predicted by both $\mathrm{Var}_t(r_{t+1})$, the variance of r_{t+1} conditional on information at t, and $\mathrm{Cov}_t(\omega_{t+1}, r_{t+1})$, the time-$t$ conditional covariance of r_{t+1} and the state variables describing the investment opportunity set, ω_{t+1}. In all of the three models, all the information required to compute $\mathrm{Cov}_t(\omega_{t+1}, r_{t+1})$ is contained in some potentially non-linear function of ω_t. We therefore test whether some non-linear function of the state variables at t can predict returns once $\mathrm{Var}_t(r_{t+1})$ is accounted for, using a MIDAS approach to estimate $\mathrm{Var}_t(r_{t+1})$. The null is that the non-linear function of ω_t cannot predict r_{t+1} once $\mathrm{Var}_t(r_{t+1})$ is accounted for.

The Bansal-Yaron state variables cannot explain the predictability of returns. We find statistically significant excess predictability (R^2 significantly greater than $\bar{R}^2$) at four out of nine horizons using annual data and six out of nine horizons using quarterly data. The MIDAS-based test also fails to reject in favour of the Bansal-Yaron state variables using either annual or quarterly data.

While there is superficially more hope for the Cecchetti-Lam-Mark model state variable, this turns out not to be robust. There is statistically significant excess predictability at only one of the nine horizons considered for the Cecchetti-Lam-Mark state variable in our main results using both annual and quarterly data. However, there are many violations in each sub-sample when we split the sample into two equal-length sub-samples, and the ability of the Cecchetti-Lam-Mark state variable to explain return predictability is not robust over time. Moreover, the MIDAS-based test fails to reject in favour of the Cecchetti-Lam-Mark state variable using either annual or quarterly data.

The only model whose state variable may explain the predictability of returns is the Campbell-Cochrane model. With annual data, the MIDAS-based test rejects in favour of the Campbell-Cochrane state variable, but there is evidence of significant excess predictability. The same is true with quarterly data. Overall, the picture is mixed, since the two state variable-based tests lead to opposite conclusions.

Apart from the question of how well these models explain the dynamics of asset returns being interesting in its own right, testing this property leads us naturally to residual-based testing. This is a standard time-series specification test, although not one that is commonly used in the context of consumption-based asset pricing models. In this setting, GMM estimation and an accompanying J-test is more common. The advantage of testing the residuals, in this case from the market return, is that it allows us to test models which are estimated in "stages" - i.e. where the estimation is not done in one single GMM implementation. Both the Campbell-Cochrane and Cecchetti-Lam-Mark models are estimated in stages in this way.

The Bansal-Yaron and Campbell-Cochrane models are two of the most prominent models designed to simultaneously explain 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 constant relative risk aversion (CRRA) 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 such as Bansal and Yaron (2004) and 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 returns in any great detail.

Huang and Zhou (2017) is the main study of how well the Bansal-Yaron and Campbell-Cochrane models explain return predictability. They develop the R^2 bound test described above, but in the context of one-step-ahead predictability of the market return with respect to several well known predictors (the book-to-market ratio, term spread, CAY, investment-to-capital ratio, new-orders-to-shipments ratio, output gap and credit expansion).[1] Huang and Zhou use Constantinides and Ghosh's (2011) inversion of the Bansal-Yaron model which renders the state variables observable. For the Campbell-Cochrane model, the state variable is unobserved and Huang and Zhou extract it as per Campbell and Cochrane's (1999) calibration. They do not estimate the model first, but condition on the extracted state variable. Huang and Zhou show that the degree of predictability in the market return is greater than can be explained by the Bansal-Yaron and Campbell-Cochrane models' state variables.

Our residual-based approach is potentially more powerful, since it can detect situations where the asset pricing model suggests too little predictability. In addition, our residual-based tests have the advantage of accounting explicitly for any initial estimation of the model or its state variables. While the Bansal-Yaron model can be inverted so that its state variables are a function of observables, this inversion is not generally possible for other asset pricing models (e.g. the Campbell-Cochrane model).

There has been little recent work on explaining own-history stock return predictability in the context of consumption-based asset pricing models. Kandel and Stambaugh (1989) propose a model with a representative CRRA investor and where consumption growth is 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 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

[1] Our adaptation is to adapt the test for q-period-ahead predictability with respect to lagged returns.

specification. They find that popular measures of serial correlation 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.

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 potentially accommodate these two puzzles. However, we also include the Cecchetti-Lam-Mark model in our results as a benchmark, since it is a model explicitly designed to explain serial correlation in returns.

Other attempts have been made to explain own-history predictability in a risk-based framework. Kim et al. (2001) proxy risk by volatility and use a volatility feedback model (where an unexpected change in volatility has an immediate impact on stock prices) with volatility following a two-state Markov-switching process. Risk adjusting returns in this way accounts for the serial correlation observed in returns. We focus on consumption-based models, which micro-found their risk factors from the start, rather than more ad hoc risk adjustments. Barroso et al. (2017) consider how conditional predictability of the short-run equity premium varies with economic and risk conditions.[2] They model the equity risk premium as a function of economic state variables.

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 consistent with the intertemporal CAPM (Barroso et al., 2017). A consumption-based asset pricing model is capable of explaining short-term conditional predictability, although no specific specification is tested.

This paper proceeds as follows. Section 2 outlines the three asset pricing models tested and their estimation. Section 3 discusses the tests we use. Section 4 briefly describes the data and reports the estimation of the asset pricing models. Section 5 presents our empirical results regarding the predictability of the model residuals and Section 6 our robustness analysis. Section 7 concludes.

[2]There are also non-risk based explanations for return predictability. These are beyond the scope of this paper.

2 The models and their estimation

2.1 Bansal-Yaron model

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

$$V_t = \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(\mathrm{E}_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{1}$$

$$x_{t+1} = \rho_x x_t + \psi_x \sigma_t \varepsilon_{t+1} \tag{2}$$

$$\Delta c_{t+1} = \mu_c + x_t + \sigma_t \eta_{t+1} \tag{3}$$

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

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

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

where V_t is the representative investor's value function, δ 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 expectation conditional on information at time t and lower-case variables denote logs of upper-case variables.

The model has three key ingredients. First, it has recursive preferences (1) à la Epstein and Zin (1989) and Weil (1989). These allow EIS and risk aversion to differ, unlike standard CRRA preferences. This is an advantage: 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).

Second, consumption growth (3) has a small predictable component (the long-run risk, x_t). Consumption news in the present affects expectations of future consumption growth, increasing the impact of current consumption news on long-run consumption and therefore the difference between present discounted values (PDVs) of dividend streams which drives returns.

Third, there is time-varying economic volatility (5) in consumption growth. This reflects time-varying economic uncertainty and is a further source of investor uncertainty and risk.

In the Bansal and Yaron (2004) calibration, the model justifies the equity premium, risk-free rate and the volatilities of the market return, risk-free rate and price-dividend ratio.

When Constantinides and Ghosh (2011) estimate the Bansal-Yaron model by GMM, the results are mixed. Simulating through the model with the estimated parameter values, the model is able to justify the market return in all specifications considered. The mean risk-free rate can be a little high, although this too is justified when the model is estimated using the identity weight matrix. Meanwhile, the J-statistic p-value is less than 0.03 in all specifications considered. However, the estimated model still generates reasonable market returns in Constantinides and Ghosh's simulations and the model may therefore still be of interest from an asset pricing point-of-view.

To estimate the model, Constantinides and Ghosh (2011) show that the log-linearised version of the Bansal-Yaron model can be inverted, allowing the unobserved state variables to be written as a linear combination of observables as follows.

$$x_t = \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 z_{m,t} \tag{6}$$
$$\sigma_t^2 = \beta_0 + \beta_1 r_{f,t} + \beta_2 z_{m,t} \tag{7}$$

where $\alpha_0, \ldots, \beta_2$ 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

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

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

In addition, they derive eight 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 and we use 15 moment conditions to allow for an overidentification test. Our set of moment conditions comprises an Euler equation for each of seven assets (the market index and six size and book-to-market double sorted portfolios, taken from Kenneth French's website), and the eight time-series restrictions.

Constantinides and Ghosh (2011) show that

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

where $r_{m,t}$ is the market return and $B_0, \ldots, B_2$ are non-linear combinations of the 12 model parameters provided in Appendix A.3. This yields a plug-in estimator of $\mathrm{E}_t\, r_{m,t+1}$, which we use as the ex-ante expected market return.

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

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

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

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)\nu_{t+1}, \tag{8}$$

where $\bar{s}$ is the steady-state s, $\bar{S} = \sigma_\nu\sqrt{\gamma/(1-\phi)}$ and $\lambda(s_t)$ is a sensitivity function given by

$$\lambda(s_t) = \begin{cases} (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1, & \text{if } s_t \leq s_{\max} \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

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

Consumption and dividends satisfy

$$\begin{aligned} \Delta c_t &= \bar{g} + \nu_t \\ \Delta d_t &= \bar{g} + w_t \end{aligned} \tag{10}$$

with Δ being the first difference operator and

$$\begin{pmatrix} \nu_t \\ w_t \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\nu^2 & \sigma_{\nu w} \\ \sigma_{\nu w} & \sigma_w^2 \end{pmatrix} \right), \tag{11}$$

where NID indicates normally and independently and identically distributed through time.

Campbell and Cochrane (1999) 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 its sample mean.[3]

In Garcia et al.'s (2004) GMM estimation of the Campbell-Cochrane model, the estimated γ is significantly greater than 0 and the δ significantly less than 1. The J-statistic p-value exceeds 0.2, although this does condition on earlier estimates of time-series parameters in the manner described below.

We estimate the Campbell-Cochrane model using a GMM procedure similar to Garcia et al. (2004). The procedure has three steps. First, we estimate the time-series parameters $\bar{g}$, σ_ν^2 and σ_w^2 in (10) by GMM. Second, we estimate α and ϕ from the linear regression

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

Based on these estimates, we generate the series s_t. We do so 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 γ of 2. This allows the series s_t to be generated as per (8) and (9).

We can then proceed to the third step: estimating the preference parameters δ and γ from the Euler equation

$$\mathrm{E}_t \left[\delta \left(\frac{S_{t+1}}{S_t} \frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + R_{t+1}) \right] - 1 = 0, \tag{12}$$

using an Euler equation for each of our seven assets. We use this new estimate of γ to generate a new s_t series, and re-estimate (12) based on this new s_t series. We iterate this procedure until the estimates of δ and γ converge. The J-statistic p-values of Garcia et al. (2004) come from their final iteration of this third step, but do not account for the initial estimation steps.

We obtain $\mathrm{E}_t\, r_{m,t+1}$ from the Campbell-Cochrane model as follows. We use the fact that $1 + R_t = (P_t + D_t)/P_{t-1}$, where P_t is the price of the asset and D_t its dividend.

[3] Campbell and Cochrane (1999) argue this is realistic as the risk-free rate varies relatively little and does not vary cyclically.

Iterating the Euler equation forwards, we have

$$P_t = \sum_{j=1}^{\infty} \delta^j \,\mathrm{E}_t \left[\left(\frac{S_{t+j}}{S_t} \frac{C_{t+j}}{C_t} \right)^{-\gamma} D_{t+j} \right] \tag{13}$$

when we impose the no-bubble condition

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

Therefore,

$$\mathrm{E}_t(1 + R_{t+1}) = \frac{\mathrm{E}_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+1+j}}{S_{t+1}} \frac{C_{t+1+j}}{C_{t+1}} \right)^{-\gamma} D_{t+1+j} + \mathrm{E}_t \, D_{t+1}}{\mathrm{E}_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+j}}{S_t} \frac{C_{t+j}}{C_t} \right)^{-\gamma} D_{t+j}}. \tag{14}$$

We estimate (14) for the market return by simulation. We simulate the series $\nu_{t+1}, \nu_{t+2}, \nu_{t+3}, \ldots$ and $w_{t+1}, w_{t+2}, w_{t+3}, \ldots$ according to (11). Based on these series, we compute the series $s_{t+1}, s_{t+2}, s_{t+3}, \ldots$, $c_{t+1}, c_{t+2}, c_{t+3}, \ldots$ and $d_{t+1}, d_{t+2}, d_{t+3}, \ldots$ conditional on s_t, c_t and d_t. We repeat this procedure 200 times, where each simulated ν_{t+1} and w_{t+1} series is of length 100. We then compute the expectation on the right-hand side of (14) as the mean of the 200 simulated realisations of the fraction inside that expectation.

2.3 Cecchetti-Lam-Mark model

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

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

Here, δ denotes the subjective discount factor and γ the coefficient of relative risk aversion. Taking (log) consumption as the appropriate endowment process,

$$\Delta c_{t+1} = \alpha_0 + \alpha_1 y_t + \varepsilon_{t+1} \,. \tag{15}$$

$y_t \in \{0, 1\}$ is a first-order Markov process and $\varepsilon_t \sim NID(0, \sigma^2)$. $y_t = 1$ denotes a bad state, so α_1 is restricted to be less than zero.

Cecchetti et al. (1990) find that, using either risk-neutral ($\gamma = 0$) or risk-averse ($\gamma = 1.7$) preferences, measures of serial correlation in the observed market return always lie within a 60% confidence interval of those measures, where the confidence intervals are generated by the model. The confidence intervals come from Monte Carlo distributions of the serial correlation statistics, obtained by simulating the model. The medians of the Monte Carlo distributions of the serial correlation statistics obtained using $\gamma = 1.7$ are closer to the observed serial correlation than the medians of the distributions using $\gamma = 0$, so Cecchetti et al. prefer the risk-averse specification. Cecchetti et al. measure serial correlation using variance ratios and Fama and French (1988) regression coefficients[4] using annual US/S&P data over 2-10 year horizons.

[4]Fama and French (1988) regression coefficients are the slope coefficient from a regression of the q-period return from t to $t + q$ on the q-period return from $t - q$ to t.

There is no guarantee that this model would simultaneously explain the equity premium and risk-free rate puzzles. Given the CRRA preferences, it probably would not. However, given the model's success in explaining market serial correlation, it is a useful benchmark for our analysis.

We use GMM to estimate δ and γ. The moment conditions comprise an Euler equation for each of our seven assets of the form

$$\mathrm{E}_t \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + R_t) \right] - 1 = 0. \tag{16}$$

We estimate the Markov switching endowment process by maximum likelihood following Hamilton (1989). In a slight deviation from Cecchetti et al. (1990), we estimate a Markov-switching process where the consumption innovation $\varepsilon_{t+1} \,|y_t \sim N(0, \sigma^2_{y_t})$, since this is more numerically stable.

$\mathrm{E}_t \, r_{m,t+1} \approx \mathrm{E}_t \left[\ln(1 + R_{m,t+1})\right]$ and Cecchetti et al. (1990) show that

$$\mathrm{E}_t \left[\ln(1 + R_{m,t+1})\right] = \mathrm{E}_t \left[\ln \left(\frac{1 + \kappa(y_{t+1})}{\kappa(y_t)} \right) + (\alpha_0 + \alpha_1 y_t) \right] \tag{17}$$

where $\kappa(y_t)$ is a non-linear function of model parameters defined in Appendix B. Since y_t is a binary variable and the conditional probabilities of each state are known, the expectation in (17) is straightforward to compute.

3 Tests

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

$$r_{m,t+1} = \mathrm{E}_t \, r_{m,t+1} + \xi_{t+1}, \tag{18}$$

where expectations are formed under the model in question and ξ_{t+1} is unforecastable at t. If the model accurately captures return dynamics, ξ_t should be MDS. If not, there is clearly something in the dynamic structure of r_t not captured by $\mathrm{E}_{t-1} \, r_t$.

We denote by θ the parameters in the model in question and define $\mathrm{E}_t \, r_{m,t+1} = \mu_{t+1}(\theta)$, to make clear the dependence of the expected returns on θ. We estimate (18) using plug-in estimators, $\mu_{t+1}(\hat{\theta})$, of $\mathrm{E}_t \, r_{t+1}$. We base our tests on the resulting residual $\xi_t(\hat{\theta})$ and denote

$$\bar{\xi} = T^{-1} \sum_{t=1}^{T} \xi_t(\hat{\theta}), \qquad \hat{s}^2 = T^{-1} \sum_{t=1}^{T} (\xi_t(\hat{\theta}) - \bar{\xi})^2.$$

We consider tests of linear and non-linear predictability in $\xi_t(\hat{\theta})$, as well as a rescaled range test. In each case, we adapt the test to cope with the fact that $\mu_t(\theta) \equiv \mathrm{E}_{t-1} \, r_{m,t}$ is estimated and this estimate, $\mu_t(\hat{\theta})$, is a function of a parameter vector estimated by GMM. It is well known that this estimation can both affect the limiting distribution of the statistics considered and induce serial dependence in the estimated residuals not present in the population.

In light of Poterba and Summers's (1988) argument that tests of the MDS null can have locally low power against certain alternatives, we use a battery of tests. Different tests have different power properties against different (local) alternatives. It therefore seems prudent to cover all bases and consider several tests. This approach bears fruit. Throughout the results, there are examples where one test fails to reject while all the others reject. It is not the case that the same test keeps failing to reject.

In addition to the MDS tests, we also use Mincer-Zarnowitz regressions (Mincer and Zarnowitz, 1969) to test the expected returns directly, and semi-parametric tests of whether the models' state variables - rather than just the functional form of the SDF - are consistent with the dynamics of the market return.

3.1 Linear predictability

A natural place to start with testing whether or not the residuals are MDS is a test based on the residuals' autocorrelations. Since the MDS null implies that all autocorrelations are zero, it makes sense to use a test statistic that incorporates autocorrelations from more than one lag. We use a weighted correlogram, of the form

$$C(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \rho(j), \tag{19}$$

where $\rho(j)$ is the j^{th} order serial correlation coefficient of ξ_t. $C(q)$ is a weighted sum of serial correlations. If $C(q) > 0$, positive autocorrelation predominates at horizon q. $C(q) < 0$ is evidence that negative autocorrelation predominates at horizon q. We consider $q \in \{2, 3, \ldots, 10\}$ years.

We use a test of the form (19) as it is a linear transformation of the variance ratio statistic. The variance ratio $VR(q)$ is the variance of the sum of q residuals divided by q times the variance of the residuals. That is $VR(q) = \text{Var}(\xi_{t+1} + \xi_{t+2} + \ldots + \xi_{t+q})/q\,\text{Var}(\xi_t)$. Since under the MDS null the residuals ξ_t and ξ_{t+j} $(j \neq 0)$ are uncorrelated, the variance ratio is equal to one under the null. Cochrane (1988) shows we can write $VR(q) = 1 + 2C(q)$, hence the connection between (19) and $VR(q)$. Poterba and Summers (1988) and Lo and MacKinlay (1989) show variance ratio tests are generally more powerful tests of the martingale difference hypothesis than unit root and autoregressive tests. The correlogram arises as a natural choice of test statistic from the Cochrane (1988) representation of the variance ratio.

In terms of estimating $C(q)$, we cannot simply treat the estimated residuals $\xi_t(\hat{\theta})$ as if they are the population residuals $\xi_t(\theta)$. The estimation of $\hat{\theta}$ affects the limiting distribution of $\hat{\rho}(j)$ under the MDS null (Delgado and Velasco, 2011). We therefore use Delgado and Velasco's (2011) transformation of the residual sample serial correlations. We denote the transformed autocorrelations by $\bar{\rho}(j)$. Delgado and Velasco start by standardising the autocorrelations so that they have a unit variance. To do this, they define the matrix A^m such that

$$(A^m)^{-1/2} \hat{\rho}^m \sim \text{N}(0, I_m),$$

with $\hat{\rho}^m = [\hat{\rho}(1), \ldots, \hat{\rho}(m)]$. To make the transformation feasible, Delgado and Velasco (2011) use Lobato et al.'s (2002) estimate of A^m

$$\hat{A}^m = \frac{1}{T\hat{s}^4} \left[g^m(0) + \sum_{j=1}^{\ell-1} \left(1 - \frac{j}{\ell}\right) \{g^m(j) + g^m(j)'\} \right]$$

where $g^m(j) = T^{-1}\sum_{t=1+j}^{T} w_t^m w_{t-j}^{m\prime}$, $w_t^m = (w_{1,t}, \ldots, w_{m,t})'$, $w_{k,t} = \left(\hat{\xi}_t(\hat{\theta}) - \bar{\xi}\right)\left(\hat{\xi}_{t-j}(\hat{\theta}) - \bar{\xi}\right)$ and ℓ is a bandwidth parameter. We use $\ell = \lceil T^{1/3} \rceil$.

Delgado and Velasco (2011) rid the estimated serial correlations collected in $\hat{\rho}^m$ of their dependence on $\hat{\theta}$ by projecting them onto the derivatives of $\hat{\xi}_t = \xi_t(\hat{\theta})$. First, define

$$\hat{\zeta}^m = \left[\hat{\zeta}(1)', \ldots, \hat{\zeta}(m)'\right]'$$

$$\hat{\zeta}(j) = \frac{1}{T\hat{s}^2}\sum_{t=j+1}^{T} \dot{\xi}_t(\hat{\theta})\left(\hat{\xi}_{t-j} - \bar{\xi}\right) + \frac{1}{T\hat{s}^2}\sum_{t=j+1}^{T} \dot{\xi}_{t-j}(\hat{\theta})\left(\hat{\xi}_t - \bar{\xi}\right)$$

$$\dot{\xi}_t(\theta) = \frac{\partial}{\partial\theta}\xi_t(\theta)$$

Then, let $\tilde{\xi}^m = \left(\hat{A}^m\right)^{-1/2}\hat{\zeta}^m$ and $\tilde{\rho}^m = \left(\hat{A}^m\right)^{-1/2}\hat{\rho}^m$. Finally, let

$$\bar{\rho}^m(j) = \frac{\check{\rho}^m(j)}{\check{s}^m(j)}$$

$$\check{\rho}^m(j) = \tilde{\rho}^m(j) - \tilde{\zeta}(j)'\left(\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\zeta}(k)'\right)^{-1}\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\rho}^m(j)$$

$$\check{s}^m(j)^2 = 1 + \tilde{\zeta}(j)'\left(\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\zeta}(k)'\right)^{-1}\sum_{k=j+1}^{m}\tilde{\zeta}(k)$$

Delgado and Velasco (2011) show that

$$\bar{\rho}^m \xrightarrow{d} N(0, I_{m-d}) \tag{20}$$

where $d = \dim(\theta)$, $\bar{\rho}^m = (\bar{\rho}(1), \ldots, \bar{\rho}(m-d))'$ and $\xrightarrow{d}$ denotes convergence in distribution. Notice that the projections sacrifice d degrees of freedom, so that only the first $m-d$ serial correlations can be transformed.

Based on (20), we estimate the weighted correlogram in (19) by

$$\bar{C}(q) = \sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)\bar{\rho}^{q-1+d}(j).$$

Because of the degrees of freedom sacrificed in the projections, we must estimate $q-1+d$ autocorrelations in order to transform the first $q-1$ autocorrelations. It follows from (20) that

$$\bar{C}(q) \xrightarrow{d} N\left(0, \left[\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)^2\right]\right)$$

under the MDS null.

3.2 Non-linear predictability

The weighted correlogram statistic is a function of the sample autocorrelations of $\hat{\xi}_t = \xi_t(\hat{\theta})$ and therefore does not exploit the full hypothesised MDS structure of $\xi_t = \xi_t(\theta)$. 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 ξ_t is MDS, the probability ξ_{t+k} is in the α quantile given ξ_t is in the α quantile should remain α. 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

$$\widehat{W}_\alpha(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \hat{\rho}_\alpha(j), \tag{21}$$

where

$$\begin{aligned}
\hat{\rho}_\alpha(j) &= \frac{\sum_{t=1}^{T-j} \psi_\alpha(\hat{\xi}_t - \hat{\mu}_\alpha)\psi_\alpha(\hat{\xi}_{t+j} - \hat{\mu}_\alpha)}{\sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\hat{\xi}_t - \hat{\mu}_\alpha)}\sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\hat{\xi}_{t+j} - \hat{\mu}_\alpha)}} \\
\psi_\alpha(\cdot) &= \alpha - \mathbb{1}(\cdot < 0) \\
\hat{\mu}_\alpha &= \operatorname*{argmin}_{m \in \mathbb{R}} \sum_{t=1}^{T} (\hat{\xi}_t - m) \times \psi_\alpha(\hat{\xi}_t - m).
\end{aligned}$$

and $\mathbb{1}(\cdot)$ is the indicator function. We evaluate (21) over the same q as in the correlograms and over a range of both extreme and moderate quantiles, namely $\alpha \in \{0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$.

We use a wild bootstrap for inference. This allows us to account for the estimation step involved in constructing $\hat{\xi}_t$. $\hat{\xi}_t$ is pre-multiplied by ι_t^* at each t, where $\mathrm{E}(\iota_t^*) = 0$ and $\mathrm{Var}(\iota_t^*) = 1$. We use Mammen's (1993) two-point distribution for ι_t^*.[5] Then, we use the bootstrapped residuals to extract a pseudo-sample of returns $r_{m,t}^*$ by the relationship

$$r_{m,t}^* = \mu_t(\hat{\theta}) + \iota_t^* \hat{\xi}_t.$$

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, $z_{m,t}^*$. We then re-estimate the asset pricing model parameters using the modified data, generating a pseudo-sample of expected returns and thus a (new) pseudo-sample of residuals.

The empirical distribution of the weighted quantilograms thus obtained is used for inference and the bootstrap procedure is repeated 200 times.[6] Notice that our procedure conditions on consumption and dividends.

3.3 Hong-Lee generalised spectral test

The Hong and Lee (2005) generalised spectral test can detect both linear and non-linear predictability. We add it to our battery of MDS tests because the known low power

[5] ι_t^* is *iid* through time and has probability mass function

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

[6] While 200 repetitions is a fairly low number, we are constrained by computational power in our ability to do more since the simulations for the Campbell-Cochrane expected returns each involve 200 repetitions themselves at each point in time in each bootstrap repetition.

problems of MDS tests (Poterba and Summers, 1988) mean it is useful to have additional tests. The test is based on the Hong (1999) generalised spectrum, corrected for the estimation of the parameters of the residual series in a way that yields a test statistic which has a nuisance parameter-free limiting distribution.

The test statistic is

$$\widehat{G}(q) = \frac{\sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right)^2 (T-j) \int_{-3}^{3} |\hat{\varsigma}_j^{(1,0)}(0,v)|^2 dW(v) - \widehat{D}(q)}{\sqrt{\widehat{E}(q)}}$$

where

$$\begin{aligned}
\widehat{D}(q) &= \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right)^2 \frac{1}{T-j} \sum_{t=j+1}^{T-1} \hat{\xi}_t^2 \int_{-3}^{3} |\hat{\pi}_{t-j}(v)|^2 dW(v) \\
\widehat{E}(q) &= 2 \sum_{j=1}^{T-2} \sum_{k=1}^{T-2} \left(1 - \frac{j}{q}\right)^2 \left(1 - \frac{k}{q}\right)^2 \int_{-3}^{3} \int_{-3}^{3} \left| \frac{1}{T - \max\{j,k\}} \right. \\
&\quad \times \left. \sum_{t=\max\{j,k\}+1}^{T} \hat{\xi}_t^2 \hat{\pi}_{t-j}(v) \hat{\pi}_{t-k}(v') \right|^2 dW(v) dW(v')
\end{aligned}$$

$W(\cdot)$ is the standard Normal distribution truncated on the interval $[-3,3]$, $\hat{\pi}(v) = e^{iv\hat{\xi}_t} - T^{-1} \sum_{t=1}^{T} e^{iv\hat{\xi}_t}$, $i = \sqrt{-1}$, and

$$\begin{aligned}
\hat{\varsigma}_j^{(1,0)}(0,v) &= \frac{\partial}{\partial u} \hat{\varsigma}_j(u,v)|_{u=0} \\
\hat{\varsigma}_j(u,v) &= \hat{\varpi}_j(u,v) - \hat{\varpi}_j(u,0)\hat{\varpi}_j(0,v) \\
\hat{\varpi}_j(u,v) &= \frac{1}{T-|j|} \sum_{t=|j|+1}^{T} e^{iu\hat{\xi}_t + iv\hat{\xi}_{t-|j|}}.
\end{aligned}$$

Under the MDS null and the technical conditions laid out in Hong and Lee (2005, p.p. 509-510), Hong and Lee show that

$$\widehat{G}(q) \xrightarrow{d} N(0,1).$$

3.4 Rescaled range

We also consider a rescaled range test. We do so as the rescaled range can be more powerful than other MDS tests in the presence of long-range dependence (Lo, 1991). The rescaled range is

$$\widehat{Q} = \frac{1}{\hat{s}\sqrt{T}} \left[\max_{k \le j \le T} \sum_{t=k}^{j} \left(\xi_t(\hat{\theta}) - \bar{\xi}\right) - \min_{k \le j \le T} \sum_{t=1}^{j} \left(\xi_t(\hat{\theta}) - \bar{\xi}\right) \right].$$

$\hat{s}^2$ is a consistent estimator of $\mathrm{Var}(\xi_t(\theta))$. Given the issue of the estimation of $\hat{\theta}$ distorting the limiting distribution of the statistic, we conduct inference using the same wild bootstrap procedure as for the quantilogram.

3.5 Mincer-Zarnowitz regressions

Aside from examining the model residuals, we can also test the forecasts themselves using a Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969). This tests (18) in a different way to the above tests. Rather than testing whether the residuals ξ_t are MDS, which is one implication of rational expectations when the models tested are correctly specified, this tests the joint null that $(a = 0) \cap (b = 1)$ in the Mincer-Zarnowitz regression

$$r_{m,t+1} = a + b \times \mu_{t+1|t}(\hat{\theta}) + e_{t+1}, \tag{22}$$

where $\mu_{t+1|t}(\hat{\theta})$ is the plug-in estimate of $\mathrm{E}_t\, r_{m,t+1}$ using information only up to time t. It is evident that $(a = 0) \cap (b = 1)$ is an implication of (18) when the model under consideration is correctly specified. If this null is rejected, the out-of-sample expected returns are systematically biased, which is inconsistent with rational expectations.

In practice, to avoid look-ahead bias, we estimate $\mu_{t+1|t}(\hat{\theta})$ using an expanding window ending at t to estimate $\hat{\theta}$. The first expanding window ends halfway through the sample, which is 1973 for the annual dataset and 1982Q1 for the quarterly dataset. This is necessary to ensure a reasonable number of observations for estimating the asset pricing model. Of course, this out-of-sample approach is potentially inefficient. It means no observations of $\mu_{t+1|t}(\hat{\theta})$ are available for the first half of the sample and the estimates of $\hat{\theta}$ are based on smaller samples than is the case in the whole-sample approach used in the correlogram, quantilogram and rescaled range tests above. Nonetheless, the Mincer-Zarnowitz approach forms a useful complement to the other tests seeing as it is based on the expected returns themselves, rather than the residuals.

3.6 Maximal predictability

Huang and Zhou (2017) develop a Wald test of whether the predictability of excess market returns, $\tilde{r}_{m,t+1} = r_{m,t+1} - r_{f,t+1}$, 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 $\mathrm{E}\,\widetilde{M}_{t+1} = 1$, being a function of a given set of state variables ω_t.[7] 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 $\mathrm{SR}(r_m)$, the market Sharpe ratio, the bound becomes

$$R^2 \leq \bar{R}^2 = \phi^2_{\omega,rf} h^2 \mathrm{SR}^2(r_m),$$

where

$$\begin{aligned}
\phi^2_{\omega,rf} &= \rho^2_{\omega,rf} \frac{\mathrm{Var}[\tilde{r}_{t+q}(\tilde{r}_t - \mu_f)]}{\mathrm{Var}(\tilde{r}_{t+q})\,\mathrm{Var}(f_t)} \\
\rho^2_{\omega,rf} &= \frac{\mathrm{Cov}[\omega_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]'\,\mathrm{Var}^{-1}(\omega_{t+q})\,\mathrm{Cov}[\omega_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]}{\mathrm{Var}[\tilde{r}_{t+q}(f_t - \mu_f)]},
\end{aligned}$$

[7] Our other tests relate to actual, not excess returns. However, $r_{f,t}$ is substantially smaller and less variable than $r_{m,t}$ and the dynamic properties of $\tilde{r}_{m,t}$ are driven by $r_{m,t}$.

and $\mu_f = \mathrm{E}(f_t)$. h is a parameter chosen by the marginal investor. We follow Cochrane and Saá-Requejo (2000) in using $h = 2$. This bound requires ω to have an elliptical distribution, which it does in all models.[8]

Huang and Zhou's (2017) test exploits the asymptotic normality of standard estimators of the mean and covariance matrix of $(r_{t+q}, f_t, r_{t+q}f_t, \omega'_{t+q})'$. These means and covariances, which comprise θ_{SR}, are all that is required to calculate the empirical R^2 and its bound. We follow Huang and Zhou and estimate θ_{SR} by GMM.

Testing whether R^2 exceeds $\bar{R}^2$ is equivalent to a one-sided test of the null $f(\theta_{SR}) \equiv R^2 - \bar{R}^2 = 0$ against the alternative that $f(\theta_{SR}) \equiv R^2 - \bar{R}^2 > 0$ (Huang and Zhou, 2017). The Wald statistic for this test is

$$W_{RA} = Tf(\hat{\theta}_{SR}) \left[\frac{df}{d\theta_{SR}} \mathrm{Var}(\hat{\theta}_{SR}) \frac{df}{d\theta_{SR}} \right]^{-1} f(\hat{\theta}_{SR}) \xrightarrow{d} \chi^2(1).$$

This procedure can then be applied to the predictive regression Fama and French (1988) use to test for serial correlation in the market return

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

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

For the Campbell-Cochrane and Cecchetti-Lam-Mark models, this test requires us to condition on our estimated state variables. The state variable for the Campbell-Cochrane model is s_t, which we extract as explained in Section 2.2. The state variable for the Cecchetti-Lam-Mark model is y_t, which we extract by estimating the Markov-switching model for consumption and taking $y_t = 1$ if the estimated smoothed probability $\Pr(y_t = 1|\mathcal{F}_{t+1}) \geq \frac{1}{2}$, where $\mathcal{F}_t$ is information available at t. The state variables for the Bansal-Yaron model are Δc_t, x_t and σ_t^2. Since we extract x_t and σ_t^2 as a linear function of $r_{f,t}$ and $z_{m,t}$, we take Δc_t, $r_{f,t}$ and $z_{m,t}$ to be the three Bansal-Yaron state variables, so that the results are not dependent on the estimation of the model.

3.7 MIDAS-based tests

A second test based of whether the models' state variables can explain the dynamics of the expected return equation comes from the Merton (1973) intertemporal CAPM (ICAPM). The ICAPM is a standard representative agent set-up where the representative investor has an increasing, concave von Neumann-Morgenstern utility function. Merton shows that, when the investment opportunity set remains constant over time, the ICAPM implies that

$$r_{m,t+1} = \pi_0 + \pi_1 \mathrm{Var}_t(r_{m,t+1}) + u_{t+1}$$

where u_{t+1} is MDS. Merton further shows that, when the investment opportunity set varies over time, the ICAPM implies that

$$r_{m,t+1} = \pi_0 + \pi_1 \mathrm{Var}_t(r_{m,t+1}) + \kappa' \mathrm{Cov}_t(r_{m,t+1}, \omega_{t+1}) + v_{t+1} \tag{24}$$

where ω_t is the vector of state variables which describe the investment opportunity set. As already discussed, the three models we consider each suggest different state variables for the investment opportunity set. For the Bansal-Yaron model, it's the risk-free interest

[8] The state variables for the Bansal-Yaron and Campbell-Cochrane models are conditionally lognormal, and the Cecchetti-Lam-Mark state variable has a binomial distribution.

rate, $r_{f,t}$ and the log market price-dividend ratio $z_{m,t}$. For the Campbell-Cochrane model, the state variable is the log surplus consumption ratio s_t, while, for the Cecchetti-Lam-Mark model, the state variable is the good/bad state indicator, y_t.

If a model's state variables are correctly specified, they will be priced. That is, $\kappa \neq 0$. A natural test of whether the model's state variables are correctly specified, then, is to test the null that $\kappa = 0$ in (24) against the two-sided alternative. In order for such a test to be feasible, a number of quantities other than the regression parameters of (24) need to be estimated. The first is $\mathrm{Var}_t(r_{m,t+1})$. We follow Ghysels et al. (2005) in using a MIDAS approach to estimate the conditional variance. We use monthly data to estimate the annual market return variance and set

$$\begin{aligned}\mathrm{Var}_t(r_{m,t+1}) &= 12\sum_{d=1}^{12} w_d(r^m_{m,t-d} - \bar{r}_{m,t})^2 \\ \bar{r}_{m,t} &= \frac{1}{12}\sum_{d=1}^{12} r^m_{m,t-d} \\ w_d &= \frac{\exp\{\pi_2 d + \pi_3 d^2\}}{\sum_{i=1}^{12}\exp\{\pi_2 i + \pi_3 i^2\}},\end{aligned}$$

where $r^m_{m,t-d}$ denotes the monthly returns in month $t-d$.

Second, we need to estimate the unobserved state variables s_t (for the Campbell-Cochrane model) and y_t (for the Cecchetti-Lam-Mark model). We do this as in the maximal predictability test and, as per the maximal predictability test, we condition on our estimates of s_t and y_t in the tests that follow.

Finally, we need to estimate $\mathrm{Cov}_t(r_{m,t+1}, \omega_{t+1})$. From the Constantinides and Ghosh (2011) inversion, we know that $(r_{f,t+1}, z_{m,t+1})$ is an AR(1) process in $(r_{f,t}, z_{m,t})$, since both x_t and σ_t^2 are separate AR(1) processes and x_t and σ_t^2 are shown to be linear in $r_{f,t}$ and $z_{m,t}$. Therefore, for the Bansal-Yaron model, we can estimate

$$r_{m,t+1} = \pi_0 + \pi_1 \mathrm{Var}_t(r_{m,t+1}) + \kappa_1 r_{f,t} + \kappa_2 z_{m,t} + v_{t+1} \tag{25}$$

and test the joint null that $\kappa_1 = \kappa_2 = 0$. We estimate (25) by quasi-maximum likelihood, following Ghysels et al. (2005), and then use an asymptotic Wald test based on a HAC residual covariance matrix estimator. We call this test the linear MIDAS test and denote the resulting test statistic $\widehat{\mathcal{W}}$.

Things are more complicated for the Campbell-Cochrane and Cecchetti-Lam-Mark models. While s_t and y_t both have the Markov property, s_{t+1} is not linear in s_t and nor is y_{t+1} linear in y_t. We must therefore use a semi-parametric approach, where (24) becomes the partially linear model

$$r_{m,t+1} = \pi_1 \mathrm{Var}_t(r_{m,t+1}) + f(\omega_t) + v_{t+1},$$

and the constant is dropped as it would not be identified. Note that the semi-parametric approach goes beyond testing the asset pricing model in question but instead tests whether its state variables are relevant. A rejection of the null that the state variables are not relevant is, therefore, evidence in favour of the model's state variables rather than the model itself, as the restrictions the model implies on the functional form of the relationship between $r_{m,t+1}$ and ω_t are not imposed.

Our test of the asset pricing model in question becomes a test of whether the $f(\omega_t)$ term has any explanatory power over $r_{m,t+1}$ once $\mathrm{Var}_t(r_{m,t+1})$ is accounted for, where

$\omega_t = s_t$ for the Campbell-Cochrane model and $\omega_t = y_t$ for the Cecchetti-Lam-Mark model. To test whether $f(\omega_t)$ term has any explanatory power over $r_{m,t+1}$ once $\text{Var}_t(r_{m,t+1})$ is accounted for, we test the null that

$$\text{E}(u_{t+1}|\omega_t) = 0 \text{ almost surely,}$$

where u_{t+1} is the regression error from (3.7). We do this using the Hsiao et al. (2007) consistent model specification test. The test statistic is given by

$$\widehat{\mathcal{I}} = \frac{1}{(T-1)(T-2)} \sum_{t=1}^{T-1} \sum_{t=1, s \neq t}^{T-1} \hat{u}_{t+1} \hat{u}_{s+1} K(\omega_t, \omega_s),$$

where K is the generalised product kernel described in Hsiao et al. (2007) and $\hat{u}_t$ is the estimated regression error from (24). As per Hsiao et al. (2007), we use the studentised version of this test, $\widehat{\mathcal{J}}$, and a wild bootstrap with a Rademacher distribution and 399 repetitions to compute the distribution under the null and p-values. We call this test the semi-parametric MIDAS test. We also compute the semi-parametric (SP) MIDAS test with $\omega_t = (r_{f,t}, z_{m,t})$ for the Bansal-Yaron state variables. This tests whether the Bansal-Yaron state variables, but not necessarily the functional form, are correct.

4 Data

Data for our main results are from the US from 1930 to 2016. The time period is annual and, as is standard in the asset pricing literature, the agent's decision interval is assumed to be the time horizon considered. We consider whether results are robust to using quarterly data and a quarterly decision interval instead as a robustness check (see Section 6.3).

The market index is the value-weighted CRSP index, obtained from WRDS. The risk-free rate is the US one-month Treasury bill, from Ibbotson Associates via French's website. The set of assets used to estimate the asset pricing models also includes the six double-sorted size/book-to-market portfolios from Ken French's website. In our robustness checks, we consider replacing the six double-sorted size/book-to-market portfolios with the five industry portfolios, also from Ken French's website, in the estimation of the models (see Section 6.2).

Consumption is seasonally adjusted per-capita non-durables and services personal consumption expenditures from the BEA. We deflate nominal data by the BEA's consumption deflator. Table 1 summarises the data.

4.1 Model estimation

Our main results relate to when the asset pricing models are estimated at the annual frequency where the set of assets used to estimate the Euler equations comprises the market return and the six double-sorted size/book-to-market portfolios. For the main results, we use the optimal weight matrix in the GMM estimation of the Campbell-Cochrane and Cecchetti-Lam-Mark models and the identity weight matrix in the GMM estimation of the Bansal-Yaron model. These specifications give the most reasonable expected returns series across the board (Section 6 gives details of the residuals for other

Table 1: Data summary statistics

	Mean	Median	Std dev	$\hat{\rho}(1)$
r_m	0.063	0.105	0.194	-0.024
r_f	0.005	0.008	0.037	0.762
Δc	0.020	0.023	0.022	0.466
Δd	0.017	0.023	0.111	0.192
z_m	3.409	3.393	0.455	0.885

Descriptive statistics for our key variables at the annual frequency over the period 1930-2016. r_m denotes the log market return, r_f the quarterly log risk-free rate (the rolled over 1 month US T-bill), Δc log consumption growth, Δd log dividend growth and z_m the log price-dividend ratio. "Std dev" denotes standard deviation and "$\hat{\rho}(1)$" estimated first-order serial correlation.

Table 2: Bansal-Yaron model estimates

μ_c	μ_d	ϕ	φ	ρ_x	ψ_x	σ	ν	σ_w	δ	ψ	γ
0.016	0.014	4.004	5.778	0.802	0.846	0.011	0.258	0.000	0.988	2.072	7.924
(0.003)	(0.013)	(0.757)	(3.661)	(0.343)	(1.214)	(0.004)	(10.17)	(0.011)	(0.208)	(40.22)	(25.41)
0.000	0.274	0.000	0.115	0.020	0.486	0.009	0.980	0.997	0.000	0.959	0.755
J-stat	43.89	p-value	0.000								

Estimates of the Bansal-Yaron model parameters using annual US data 1930-2016. 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.

specifications; because actual returns are the sum of the expected return and the residual, only models with reasonable residual series will have reasonable expected returns).

Many of the other specifications do not give reasonable expected returns series. This is less surprising than it might seem given the challenges of identifying asset pricing models using GMM, as emphasised by Cheng et al. (2022). We look only at specifications where the expected returns are plausible. As much as our focus is on the dynamics of returns, rather than the levels, the first and second moments are related. Serial correlation (a centred second moment) depends on the first moment. But, even if we only used uncentred second moments, there is no reason to think that a model that fails to fit the first moment would fit the second. Moreover, if it did, it would be of little practical relevance for pricing assets. While we focus on the specification that generally gives the most reasonable expected returns, our results are robust to considering other specifications giving reasonable expected returns.

Table 2 suggests the Bansal-Yaron model may be mis-specified. The J-statistic has a vanishingly small p-value. However, the estimated risk aversion is positive and the estimated time discount factor is less than one. Table 5 shows the Bansal-Yaron model has the highest absolute mean and median residuals.

Table 3 shows the estimated Campbell-Cochrane model parameters. $\hat{\gamma}$ is positive but not significantly so, although the subjective discount factor is significantly less than one. The J-test rejects the model's Euler equations. Nonetheless, this is only indicative of how well specified the Euler equations are. The Euler equation estimation conditions on earlier estimates of time-series parameters ($\bar{g}$, $\mathrm{Var}(\Delta c)$, $\mathrm{Var}(\Delta d)$, $\mathrm{Cov}(\Delta c, \Delta d)$, α and ϕ), yet the over-identification test in the third panel of Table 3 does not account for this estimation. We cannot firmly reject the model on this basis. Table 5 shows that the mean residual is close to zero, just 0.7%. The Campbell-Cochrane model therefore seems

Table 3: Campbell-Cochrane model estimates

$\bar{g}$	$\text{Var}(\Delta c)$	$\text{Var}(\Delta d)$	$\text{Cov}(\Delta c, \Delta d)$	α	ϕ	δ	γ
0.021	4.06×10^{-4}	0.012	0.001	0.424	0.879	0.929	0.043
(0.003)	(1.52×10^{-4})	(0.004)	(0.001)	(0.173)	(0.050)	(0.020)	(0.836)
0.000	0.007	0.003	0.056	0.017	0.000	0.000	0.959
J-stat	0.089			R^2	0.783	J-stat	36.24
p-value	0.766					p-value	0.000

Estimates of the Campbell-Cochrane model parameters using annual US data 1930-2016. Each panel (set of columns) refers to a separate estimation. The estimates of δ and γ, and the associated p-values, 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.

Table 4: Cecchetti-Lam-Mark model estimates

(a) Consumption model

α_0	α_1	p	q	σ_0^2	σ_1^2
0.023	-0.016	0.956	0.876	0.012	0.040

(b) Preference parameters

δ	γ
0.966	2.431
(0.290)	(15.38)
0.001	0.874
J-stat	37.18
p-value	0.000

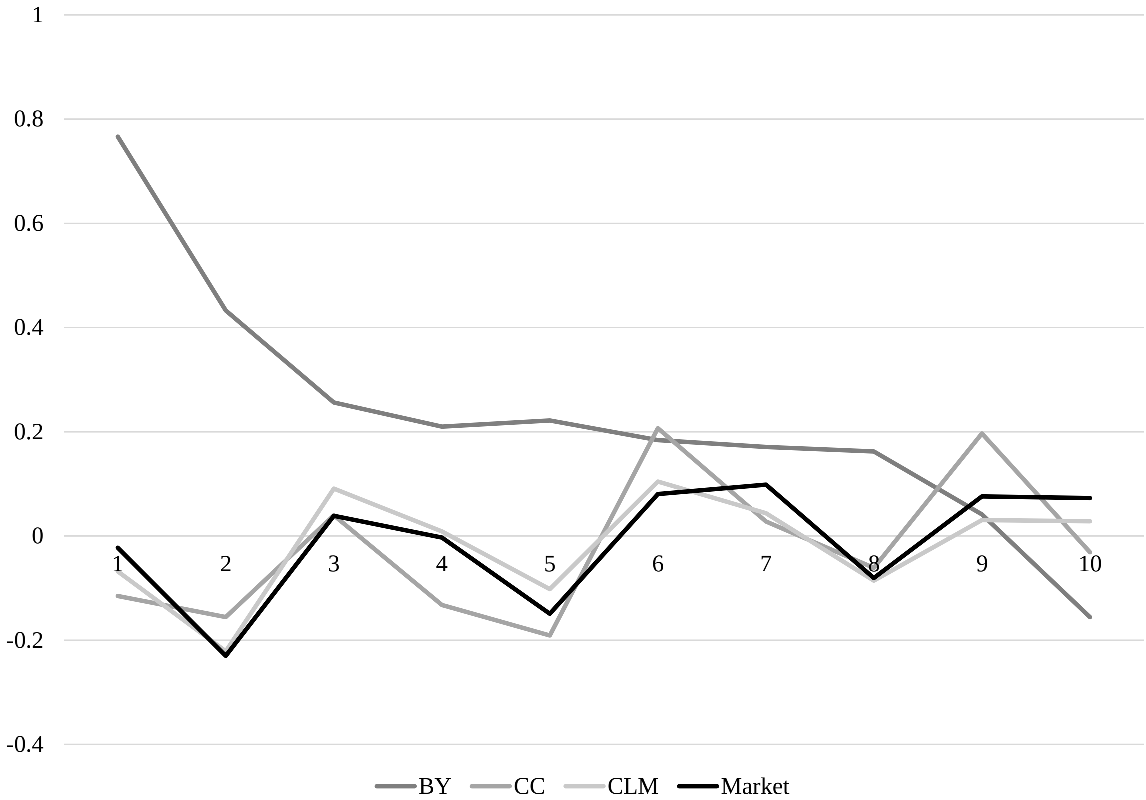
Estimates of the Cecchetti-Lam-Mark model parameters, estimated using annual US data 1930-2016. Panel (a) presents point estimates only. In panel (b), point estimates are displayed in the first row, standard errors (in parentheses) in the second and p-values in the third. All p-values are asymptotic.

Table 5: Properties of $\hat{\xi}_t$

Model	Mean	Median	Std dev	$\hat{\rho}(1)$
Bansal-Yaron	-0.066	-0.097	0.444	0.589
Campbell-Cochrane	0.007	-0.012	0.210	-0.115
Cecchetti-Lam-Mark	-0.017	0.014	0.191	-0.080

Summary statistics for the model-implied ex-ante residuals. "Std dev" denotes standard deviation and "$\hat{\rho}(1)$" estimated first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

Figure 1: Market and model autocorrelation functions



Autocorrelation functions for the market return and the model-implied ex-ante expected returns. Serial correlation is computed up to lag 10. The models are estimated and expected returns computed over 1930-2016. These estimates of the model-implied autocorrelation functions are biased due to the estimation of the parameters of the expected returns and it is therefore difficult to draw many firm conclusions from this figure, which is provided for illustrative purposes only.

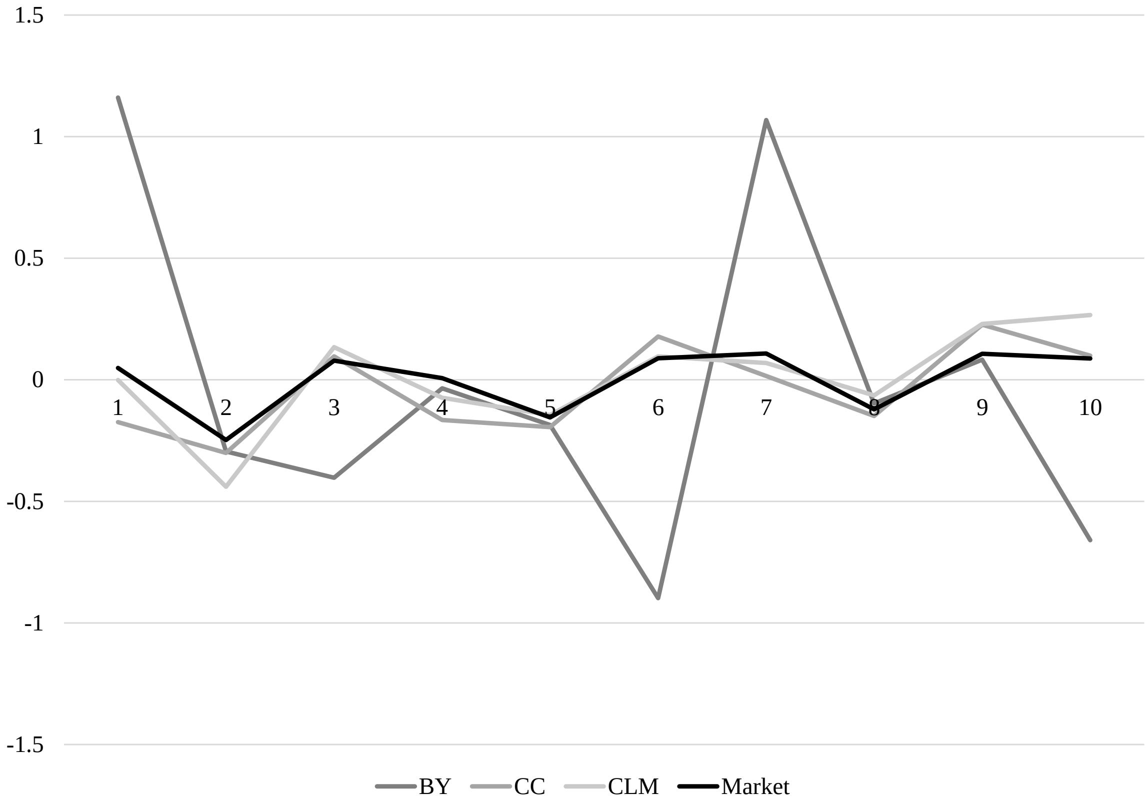
to give reasonable expected returns, despite the J-test.

Table 4 shows that the Cecchetti-Lam-Mark model preference parameter estimates are also generally reasonable. The subjective discount factor is less than one and the utility curvature greater than zero, albeit not significantly so. The Euler equations are rejected by the J-test, but this test does not enforce the Markov-switching structure on consumption growth. Enforcing this structure may still yield reasonable expected returns. Table 5 suggests this is indeed the case. The mean residual for the Cecchetti-Lam-Mark model is fairly low at around -1.7% a year.

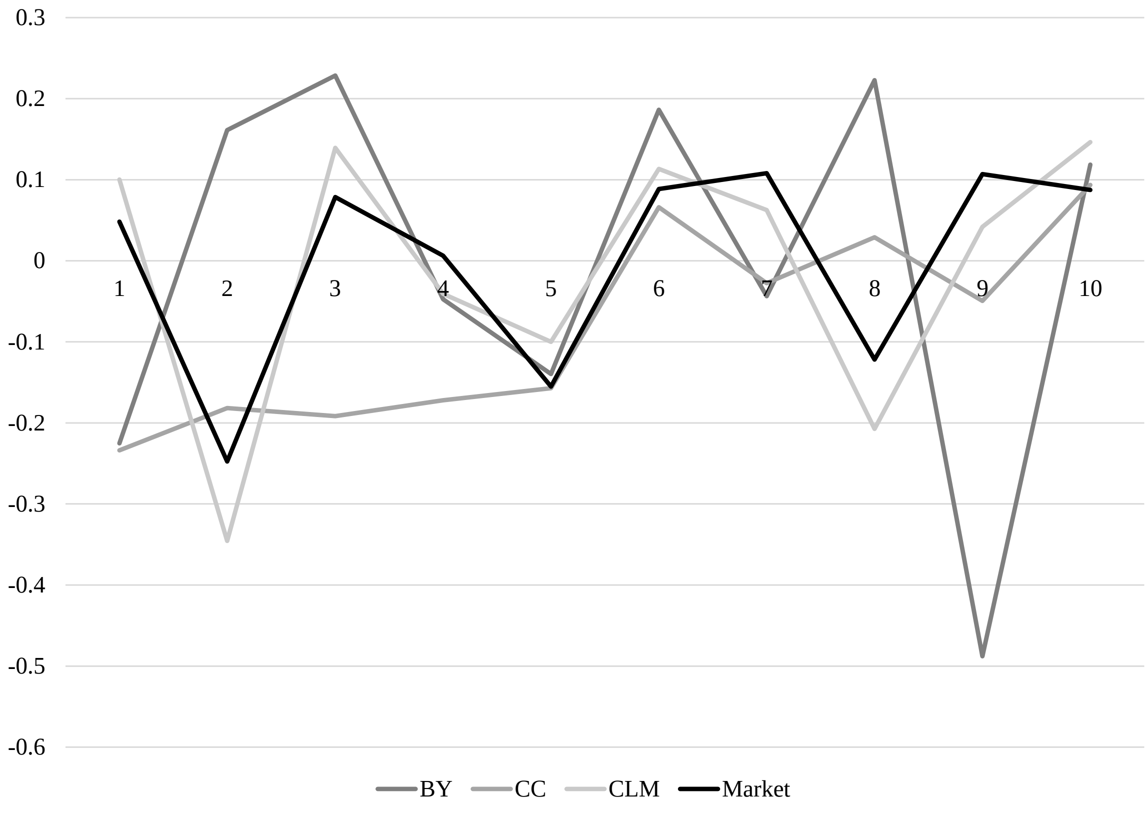
Figure 1 shows the autocorrelation functions of the observed market return and the model-implied ex-ante expected returns. This graph is only indicative. We must be mindful of the distortions in the model-implied autocorrelation functions induced by parameter estimation. In the graph, the Bansal-Yaron is a long way from matching the market autocorrelation function. The Campbell-Cochrane and Cecchetti-Lam-Mark model expected return autocorrelations are fairly close to the observed market autocorrelations.

To remove the effect of estimation in the autocorrelations of the expected returns, we can apply the Delgado and Velasco (2011) procedure to them. Note that the Delgado and Velasco procedure transforms the standardised autocorrelations $\tilde{\rho}^m = (\hat{A}^m)^{1/2}\hat{\rho}^m$ to give the transformed standardised autocorrelations $\bar{\rho}^m$. So, in order to see the effect of the transformation, we need to consider the untransformed standardised autocorrelations and the transformed standardised autocorrelations. These are shown in panel (a) of Figure

Figure 2: Market and model standardised autocorrelation functions



(a) Untransformed standardised autocorrelation function ($\tilde{\rho}$)



(b) Transformed standardised autocorrelation function ($\bar{\rho}$)

Untransformed and transformed standardised (divided by standard error) autocorrelation for the model-implied ex-ante expected returns compared to (untransformed) standardised autocorrelation for the market. Serial correlation is computed up to lag 10.

2, where $m = 10 + \dim(\theta)$ for each model. Panel (b) shows the transformed standardised autocorrelations, $\bar{\rho}^m$. The market autocorrelation in panel (b) remains $\tilde{\rho}^m$, since there is no adjustment needed.

We can see that, both with and without the Delgado and Velasco (2011) adjustment, the Bansal-Yaron model's standardised autocorrelations have some large deviations from those of the market. Note that Figure 2 shows autocorrelations which have been divided by their standard errors, which is why some of the standardised autocorrelations for the Bansal-Yaron model are greater than one in panel (a). Oddly, the Campbell-Cochrane standardised autocorrelations appear to be closer to those of the market before applying the adjustment. This would imply that the bias in the autocorrelation function of the Campbell-Cochrane expected returns arising from the estimation of the model parameters was making the Campbell-Cochrane autocorrelations artificially close to the market's autocorrelations. The adjustment does not appear to impact how close the Cecchetti-Lam-Mark autocorrelations are to the market autocorrelations: they seem to be close in both cases.

5 Serial dependence in the model residuals

Our results for the Bansal-Yaron model are in Table 6. We clearly reject the null that the Bansal-Yaron residuals are MDS, with the null being rejected by the rescaled range test and by the majority of the weighted quantilograms. 59 out of 81 quantilogram tests reject the MDS null at the 10% level and 58 of them reject the MDS null at the 5% level. Neither the weighted correlogram nor the Hong-Lee test reject the MDS null at any lag, which shows the value of not relying on just one test. In addition, the Mincer-Zarnowitz F-test rejects the Bansal-Yaron model and the linear MIDAS test fails to reject the null that the Bansal-Yaron state variables are not relevant for the expected return, conditioning on $\mathrm{Var}_t(r_{m,t+1})$.

The semi-parametric MIDAS test does not reject the null that the Bansal-Yaron state variables are not relevant for the expected return, conditioning on $\mathrm{Var}_t(r_{m,t+1})$, too. Moreover, the maximal predictability results suggest that the Bansal-Yaron state variables do not explain observed predictability, either. Changing the functional form of the SDF would not enable a model based on the Bansal-Yaron state variables to explain the dynamics of returns. There are extremely significant exceedences of the R^2 bound, $\bar{R}^2$, at four horizons: four, five, six and seven years.

However, we express some caution regarding these maximal predictability results for two reasons. First, $\bar{R}^2$ 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 Bansal-Yaron state variables or no predictability is consistent with these state variables. Second, the parameters of R^2 and $\bar{R}^2$ are jointly estimated using GMM. The R^2 does not come directly from a regression themselves. The methods ought to be equivalent but it is not computationally possible to satisfy the moment conditions exactly here, despite the system 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 the Bansal-Yaron model as it is for the Campbell-Cochrane and Cecchetti-Lam-Mark models, even though it should be. These discrepancies highlight the numerical challenges of the GMM estimation undertaken to compute the tests. However, these numerical issues do not affect the

Table 6: Bansal-Yaron model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.015	-0.083	-0.039	-0.013	-0.032	-0.048	-0.095	-0.083	-0.157
(Std Err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.783	0.303	0.702	0.914	0.808	0.745	0.551	0.628	0.390

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.013	-0.016	-0.020	-0.024	-0.028	-0.032
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.05	0.219	0.277	0.335	0.456	0.584	0.667	0.724	0.763	0.788
	0.81	0.38	0.31	0.27	0.23	0.23	0.21	0.19	0.14
0.1	0.145	0.245	0.367	0.471	0.542	0.609	0.664	0.695	0.735
	0.05	0.02	0.01	0.01	0.01	0.01	0.00	0.00	0.00
0.25	0.208	0.337	0.485	0.635	0.741	0.829	0.910	0.967	1.009
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.5	0.185	0.318	0.443	0.558	0.665	0.745	0.830	0.898	0.947
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.75	0.224	0.412	0.562	0.707	0.817	0.902	0.972	1.030	1.084
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	0.119	0.181	0.261	0.349	0.410	0.464	0.498	0.526	0.549
	0.15	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.157	0.236	0.328	0.376	0.403	0.416	0.421	0.421	0.417
	0.63	0.30	0.14	0.08	0.04	0.03	0.02	0.01	0.01
0.99	0.195	0.255	0.284	0.300	0.310	0.316	0.320	0.323	0.324
	0.13	0.31	0.50	0.65	0.79	0.91	1.00	0.91	0.85

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	1.476	1.482	1.489	1.496	1.503	1.510	1.517	1.524	1.531
p-value	0.140	0.138	0.136	0.135	0.133	0.131	0.129	0.127	0.126

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.093	0.079	145.3
(Std err)	(0.019)	(0.090)	
p-value	0.000	0.386	0.000

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one.

Table 6: Bansal-Yaron model results

(e) Rescaled range UPDATED

$\widehat{Q}$	1.566
p-value	0.00

(f) Linear MIDAS test

$\widehat{\mathcal{W}}$	2.593
p-value	0.273

(g) SP MIDAS test

$\widehat{\mathcal{J}}$	-0.573
p-value	0.190

(h) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.100	0.044	0.462	0.002	0.038	0.043	0.125	0.028	0.000
$\bar{R}^2$	21.68	0.122	-12.29	-59.18	-0.725	-2.738	1993	67.57	1.058
Wald stat	-	-	42.19	54.06	50.13	5.156	-	-	-
p-value	-	-	0.000	0.000	0.000	0.023	-	-	-

$\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath. $\widehat{\mathcal{W}}$ denotes the MIDAS Wald statistic and its asymptotic p-value is given beneath. $\widehat{\mathcal{J}}$ is the estimated Hsiao et al. (2007) consistent model specification statistic from the semi-parametric MIDAS model. Its bootstrapped p-value is given beneath. Panel (h) reports tests of the null that the market return is no more predictable than implied by the Bansal-Yaron model state variables (i.e. $R^2 \leq \bar{R}^2$). The Wald statistic and its asymptotic p-value are reported.

maximal predictability tests for the Campbell-Cochrane or Cecchetti-Lam-Mark models so may simply be a further reflection of the mis-specification of the Bansal-Yaron state variables. Overall, the best available evidence is that the state variables of the Bansal-Yaron model cannot explain the predictability of market returns.

Our main results regarding the Campbell-Cochrane model are in Table 7. We reject the null that the Campbell-Cochrane residuals are MDS: the correlogram rejects the MDS null at all lags considered. However, the Hong-Lee test and rescaled range provide no rejections and only two of the 81 quantilograms reject the MDS null at the 10% level. The Mincer-Zarnowitz F-test failed to reject the model as well, even though the estimated correlation between expected and actual returns was negative, due to large standard errors. This again shows the value of not relying on only one test statistic.

Turning to our state variable tests, the semi-parametric MIDAS test rejects, at the 10% level, the null that the Campbell-Cochrane state variables are irrelevant for expected returns once the conditional return variance is accounted for. However, there are three significant exceedences of the R^2 bound in the maximal predictability test. It therefore appears possible that models based on the surplus consumption state variable may be able to explain the dynamics of returns, although the evidence is mixed, since the two state variable tests point in opposite directions.

Table 8 shows the results for the Cecchetti-Lam-Mark model. The residuals are clearly not MDS. The correlogram rejects the MDS null from $q = 5$ onwards and the rescaled range also rejects the MDS null. Both rejections suggest negative serial dependence: that higher values are followed by lower ones. Neither the quantilogram nor the Hong-Lee tests provide any rejections of the MDS null. This serves to further illustrate the power issues of MDS tests and justify our approach of considering multiple different tests. Moreover, like with the Campbell-Cochrane model, the Mincer-Zarnowitz F-test fails to reject the Cecchetti-Lam-Mark model despite there being a negative correlation between expected and actual returns due to large standard errors.

The semi-parametric MIDAS test does not reject the null that the Cecchetti-Lam-

Table 7: Campbell-Cochrane model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.117	-0.170	-0.196	-0.239	-0.304	-0.341	-0.430	-0.442	-0.452
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.030	0.035	0.052	0.043	0.022	0.020	0.007	0.010	0.013

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.014	0.017	0.026	-0.006	-0.059	-0.064	-0.030	-0.009
	0.06	0.61	0.57	0.78	0.69	0.70	0.89	0.77	0.06
0.05	-0.006	-0.028	-0.012	-0.020	-0.080	-0.120	-0.088	0.024	-0.018
	0.23	0.97	0.74	0.57	0.71	0.64	0.73	0.83	0.23
0.1	-0.009	-0.043	-0.052	-0.086	-0.114	-0.132	-0.129	0.038	-0.026
	0.56	0.98	0.82	0.71	0.76	0.79	0.80	0.83	0.25
0.25	-0.013	-0.058	-0.070	-0.140	-0.150	-0.141	-0.137	0.037	-0.035
	0.70	0.96	0.84	0.69	0.93	0.98	0.97	0.95	0.25
0.5	-0.016	-0.073	-0.101	-0.183	-0.179	-0.157	-0.163	0.029	-0.044
	0.86	0.92	0.96	0.83	0.99	0.90	0.80	0.92	0.25
0.75	-0.020	-0.042	-0.099	-0.217	-0.186	-0.150	-0.160	0.015	-0.053
	0.91	0.94	0.99	0.90	0.92	0.75	0.75	0.94	0.36
0.9	-0.024	-0.001	-0.092	-0.234	-0.190	-0.141	-0.173	-0.001	-0.062
	1.00	0.94	0.96	0.97	0.88	0.66	0.70	1.00	0.38
0.95	-0.028	0.029	-0.098	-0.230	-0.201	-0.143	-0.197	-0.020	-0.071
	0.94	0.96	0.93	0.96	0.82	0.62	0.67	0.99	0.44
0.99	0.004	0.052	-0.113	-0.212	-0.198	-0.142	-0.222	-0.040	-0.080
	0.99	0.86	0.81	0.85	0.62	0.52	0.48	0.61	0.38

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.381	0.386	0.388	0.386	0.383	0.379	0.373	0.365	0.345
p-value	0.703	0.699	0.698	0.699	0.702	0.705	0.709	0.715	0.730

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.080	–0.287	1.342
(Std err)	(0.050)	(1.070)	
p-value	0.117	0.790	0.273

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one.

Table 7: Campbell-Cochrane model results

(e) Rescaled range

$\widehat{Q}$	0.911
p-value	0.28

(f) SP MIDAS test

$\widehat{\mathcal{J}}$	-0.466
p-value	0.063

(g) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.057	0.022	0.012	3.4×10^{-4}	0.029	0.035	0.039	0.020	0.002
$\bar{R}^2$	0.127	0.001	0.067	1.29×10^{-5}	0.010	0.045	0.012	0.025	0.030
Wald stat	-	16.91	-	0.185	7.920	-	4.618	-	-
p-value	-	0.000	-	0.667	0.005	-	0.032	-	-

$\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath. $\widehat{\mathcal{J}}$ is the estimated Hsiao et al. (2007) consistent model specification statistic from the semi-parametric MIDAS model. Its bootstrapped p-value is given beneath. Panel (g) reports tests of the null that the market return is no more predictable than implied by the Campbell-Cochrane model state variables (i.e. $R^2 \leq \bar{R}^2$). The Wald statistic and its asymptotic p-value are reported.

Mark state variable is not relevant for expected returns once $\mathrm{Var}_t(r_{m,t+1})$ is accounted for. However, there is only one significant exceedence of the R^2 bound in the maximal predictability tests, at $q = 2$. This apparent conflict is resolved using quarterly data, where the semi-parametric MIDAS test continues to fail to reject in favour of the Cecchetti-Lam-Mark state variable and the maximal predictability tests do reject the Cecchetti-Lam-Mark state variable (see Section 6.3).

6 Robustness

We consider the robustness of our results to (i) using the identity weight matrix in GMM estimation rather than the optimal weight matrix, (ii) using the five Fama-French industry portfolios in place of the six Fama-French size/value portfolios when estimating the asset pricing models and (iii) using quarterly data instead of annual data. Overall, we find that, where the models produce reasonable residual and expected returns series, they cannot explain return dynamics, either in the MDS tests on the residuals or Mincer-Zarnowitz tests on expected returns.

In terms of the state variable tests, the finding that the maximal predictability tests reject the Bansal-Yaron state variables is robust to using quarterly data. However, the semi-parametric MIDAS test suggests more promise for the Bansal-Yaron state variables. That the Campbell-Cochrane state variables cannot explain expected returns conditional on $\mathrm{Var}_t(r_{m,t+1})$ is a finding robust to using quarterly data. The finding that the Cecchetti-Lam-Mark model may be able to explain the predictability of returns survives switching to quarterly data in the whole sample, but this finding is not robust over time. When we split the sample period into two equal-length sub-samples, we get many more significant R^2 bound exceedences in both sub-samples than in the whole sample. The failure of the semi-parametric MIDAS test to reject in favour of the Cecchetti-Lam-Mark state variable is completely robust.

We consider the robustness of the residual-based tests (i.e. the correlogram, quantilogram, Hong-Lee tests and rescaled range) only in scenarios where the model provides

Table 8: Cecchetti-Lam-Mark model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	0.043	-0.087	-0.201	-0.275	-0.342	-0.361	-0.428	-0.455	-0.526
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.425	0.278	0.047	0.020	0.010	0.014	0.007	0.008	0.004

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.009	0.018	0.080	0.165	0.262	0.359	0.448	0.526	0.587
	0.39	0.77	0.98	0.78	0.74	0.67	0.67	0.66	0.71
0.05	-0.006	-0.017	-0.036	-0.053	-0.070	-0.090	-0.104	-0.116	-0.126
	0.72	0.78	0.72	0.66	0.61	0.57	0.57	0.58	0.47
0.1	-0.004	-0.014	-0.030	-0.047	-0.063	-0.084	-0.099	-0.116	-0.132
	0.76	0.67	0.62	0.60	0.53	0.51	0.47	0.45	0.39
0.25	0.015	0.004	-0.005	-0.016	-0.026	-0.039	-0.053	-0.072	-0.094
	0.86	0.75	0.70	0.67	0.60	0.57	0.53	0.48	0.42
0.5	0.031	0.028	0.032	0.035	0.029	0.016	0.000	-0.021	-0.045
	0.99	0.95	0.87	0.85	0.78	0.73	0.72	0.70	0.59
0.75	0.089	0.115	0.147	0.170	0.179	0.178	0.167	0.145	0.120
	0.97	0.87	0.87	0.85	0.81	0.77	0.77	0.73	0.66
0.9	0.069	0.081	0.082	0.082	0.076	0.063	0.049	0.030	0.009
	0.98	0.92	0.88	0.86	0.86	0.80	0.76	0.73	0.60
0.95	0.009	0.009	-0.004	-0.018	-0.035	-0.055	-0.075	-0.098	-0.119
	0.91	0.84	0.75	0.74	0.69	0.67	0.64	0.63	0.49
0.99	-0.009	-0.025	-0.051	-0.100	-0.170	-0.259	-0.367	-0.495	-0.650
	0.41	0.45	0.47	0.49	0.52	0.51	0.52	0.53	0.48

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.884	0.897	0.904	0.907	0.907	0.903	0.894	0.882	0.867
p-value	0.377	0.370	0.366	0.364	0.364	0.367	0.371	0.378	0.386

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.079	–0.154	0.763
(Std err)	(0.072)	(0.940)	
p-value	0.278	0.871	0.473

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one.

Table 8: Cecchetti-Lam-Mark model results

(e) Rescaled range

$\widehat{Q}$	0.698
p-value	0.02

(f) SP MIDAS test

$\widehat{\mathcal{J}}$	-0.725
p-value	0.602

(g) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.057	0.022	0.012	0.000	0.029	0.035	0.039	0.020	0.002
$\bar{R}^2$	0.006	0.026	0.033	0.160	0.194	0.026	0.050	0.362	0.346
Wald stat	134.1	-	-	-	-	0.189	-	-	-
p-value	0.000	-	-	-	-	0.664	-	-	-

$\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath. $\widehat{\mathcal{J}}$ is the estimated Hsiao et al. (2007) consistent model specification statistic from the semi-parametric MIDAS model. Its bootstrapped p-value is given beneath. Panel (g) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$). The Wald statistic and its asymptotic p-value are reported.

Table 9: Properties of $\hat{\xi}_t$ - Identity matrix

Model	Mean	Median	Std dev	$\hat{\rho}(1)$
Campbell-Cochrane	-0.466	-0.397	0.289	0.455
Cecchetti-Lam-Mark	-0.038	-0.040	0.191	-0.088

Summary statistics for the model-implied ex-ante residuals. “Std dev” denotes standard deviation and “$\hat{\rho}(1)$” estimated first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

credible residuals, and therefore credible expected returns. There is no point checking the second moment of a model that fits poorly in terms of the first moment, as one would not use it to price assets anyway. Moreover, the centred second moment (e.g. serial correlation coefficient) is a function of the first moment.

For the robustness of the state variable tests, note that the state variables in the Bansal-Yaron and Cecchetti-Lam-Mark are independent of the asset sets or GMM weighting matrices used. As such, the maximal predictability results for these models depend only on the data frequency and sample period. The extraction of the Campbell-Cochrane state variable depends on, amongst other things, the estimated utility curvature. Therefore, the (estimated) state variable does depend on the asset set and GMM weighting matrix. As a result, we consider the robustness of the Campbell-Cochrane maximal predictability tests in each of the scenarios set out above.

6.1 Identity weight matrix

Table 9 shows that only the Cecchetti-Lam-Mark model gives rise to a credible expected returns series: the mean residual of 3.8% implies a mean expected market return of 10% a year. The Campbell-Cochrane model's average residual of -46.6% coupled with the mean market return of 6.3% implies a mean expected market return of over 50% a year under the Campbell-Cochrane model. This is almost nine times the actual value, and the expected returns do not form a credible financial time series. As noted earlier, the main results for the Bansal-Yaron model already use the identity weight matrix since the estimates using an optimal weight matrix do not converge.

The results of the MDS tests for the Cecchetti-Lam-Mark residuals when the model is estimated with the identity weight matrix are shown in Table 10. They paint a similar picture to the results with the optimal weight matrix: the correlograms reject the MDS null (at the 10% level) from $q = 4$ onwards and the rescaled range rejects the MDS null too. Again, both tests imply anti-persistence in the residuals, while the quantilogram and Hong-Lee tests do not reject the null. In addition, and in contrast to the results with the optimal weight matrix, the Mincer-Zarnowitz F-test now rejects the model.

Notice that the choice of weight matrix does not affect the extraction of the Bansal-Yaron or Cecchetti-Lam-Mark state variables, so these MIDAS and maximal predictability test results are unchanged. The GMM estimation for R^2 and $\bar{R}^2$ using the extracted Campbell-Cochrane state variable did not converge, so maximal predictability results are not available. The semi-parametric MIDAS test is available for the Campbell-Cochrane and now gives a slightly stronger rejection of in favour of the Campbell-Cochrane state variable ($\widehat{\mathcal{J}} = -0.466$, p-value $= 0.040$).

6.2 Industry portfolios

Table 11 shows summary statistics of the residuals where we replace the six Fama-French size/value portfolios with the five Fama-French industry portfolios in the set of assets used to estimate the asset pricing models. As when using the size/value portfolios, GMM estimation of the Bansal-Yaron model does not converge when using the optimal weight matrix. Only the Campbell-Cochrane model estimated with the identity weight matrix produces a credible residual, and therefore expected return, series. With a mean residual of -11.5% and a mean market return of 6.3%, the mean expected market return is 17.8%. Even this is stretching the bounds of credibility. But there is little harm in considering the robustness of the residual-based tests in this scenario in any case.

The Campbell-Cochrane model results when estimating the model using the industry portfolios and the identity weight matrix are shown in Table 12. We resoundingly reject the null that the residuals are MDS. The correlogram test produces four rejections at the 10% level, at the two shortest and two longest horizons considered. There are 72 rejections of the MDS null out of 81 quantilogram tests. The 99th percentile is the only one where we do not reject the MDS null. This is very different to the results when using size/value portfolios and optimal weight matrix, where there are only two such rejections. Similarly, the Mincer-Zarnowitz F-statistic now rejects the model, whereas it did not when using the size/value portfolios and optimal weight matrix. While the Hong-Lee test produces no rejections, the rescaled range test also rejects the MDS null. Whether or not one considers the residuals to be a plausible financial time series, they are not MDS and the model is again rejected.

Turning to the state variable tests, note again that the Bansal-Yaron and (extracted) Cecchetti-Lam-Mark state variables are unaffected by the change in the assets set, as well as the change in weight matrix. The Campbell-Cochrane state variable is, however, affected. The semi-parametric MIDAS results are very similar to when the size/value portfolios are used, with the tests again rejecting in favour of the Campbell-Cochrane state variable. When using the industry portfolios and the optimal weight matrix, we obtain a semi-parametric MIDAS test statistic of $\widehat{\mathcal{J}} = -0.466$ (p-value $= 0.058$), while with the identity weight matrix we obtain $\widehat{\mathcal{J}} = -0.466$ (p-value $= 0.053$). The maximal

Table 10: Cecchetti-Lam-Mark model results - Identity matrix

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	0.026	-0.122	-0.190	-0.260	-0.300	-0.311	-0.351	-0.381	-0.390
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.633	0.129	0.059	0.028	0.025	0.034	0.028	0.026	0.032

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.013	-0.016	-0.020	-0.024	-0.028	0.004
	0.40	0.75	0.99	0.83	0.80	0.70	0.70	0.69	0.76
0.05	-0.014	-0.028	-0.043	-0.058	-0.073	-0.042	-0.022	-0.007	0.004
	0.74	0.80	0.78	0.70	0.69	0.65	0.61	0.59	0.47
0.1	0.017	0.011	-0.018	-0.003	0.012	0.047	0.060	0.067	0.063
	0.68	0.66	0.54	0.55	0.51	0.49	0.46	0.45	0.38
0.25	0.026	-0.053	-0.111	-0.111	-0.136	-0.157	-0.154	-0.146	-0.147
	0.74	0.70	0.60	0.54	0.53	0.49	0.47	0.41	0.37
0.5	-0.030	-0.136	-0.162	-0.192	-0.260	-0.301	-0.318	-0.337	-0.350
	0.70	0.66	0.60	0.48	0.46	0.44	0.35	0.33	0.25
0.75	-0.059	-0.100	-0.140	-0.134	-0.140	-0.119	-0.126	-0.130	-0.137
	0.64	0.62	0.56	0.50	0.46	0.47	0.44	0.39	0.30
0.9	-0.003	0.034	0.084	0.139	0.199	0.227	0.251	0.275	0.310
	0.76	0.79	0.65	0.57	0.49	0.44	0.42	0.36	0.26
0.95	-0.030	0.024	0.070	0.137	0.197	0.234	0.258	0.272	0.280
	0.79	0.68	0.58	0.56	0.49	0.45	0.40	0.38	0.33
0.99	-0.009	-0.018	-0.024	-0.030	-0.036	-0.041	-0.047	-0.052	-0.057
	0.31	0.38	0.40	0.35	0.37	0.39	0.38	0.40	0.34

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.782	0.795	0.801	0.803	0.803	0.800	0.793	0.783	0.770
p-value	0.434	0.427	0.423	0.422	0.422	0.424	0.428	0.434	0.441

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.132	–0.459	4.039
(Std err)	(0.231)	(1.665)	
p-value	0.571	0.784	0.025

(e) Rescaled range

$\widehat{Q}$	0.694
p-value	0.01

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one. In Panel (e), $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 11: Properties of $\hat{\xi}_t$ - Industry portfolios

Model	Mean	Median	Std dev	$\hat{\rho}(1)$
Optimal weight matrix				
Bansal-Yaron	-	-	-	-
Campbell-Cochrane	-0.250	-0.279	0.232	-0.045
Cecchetti-Lam-Mark	-0.185	-0.158	0.194	-0.041
Identity weight matrix				
Bansal-Yaron	-0.685	-0.715	0.426	0.564
Campbell-Cochrane	-0.115	-0.130	0.257	-0.126
Cecchetti-Lam-Mark	-0.242	-0.098	0.376	0.567

Summary statistics for the model-implied ex-ante residuals. "Std dev" denotes standard deviation and "$\hat{\rho}(1)$" estimated first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

predictability test results are also very similar.

The GMM estimation of R^2 and $\bar{R}^2$ does not converge for the Campbell-Cochrane state variable extracted based on parameter estimates using the optimal weight matrix to estimate the model. The estimation does converge, though, when the identity weight matrix is used in the Campbell-Cochrane model estimation. These untabulated results provide three significant exceedences of the R^2 bound, all at the same horizons as when using the size/value portfolios and optimal weight matrix to estimate the Campbell-Cochrane model and extract the surplus consumption state variable.

6.3 Quarterly data

Returning to using the six size/value portfolios in the set of assets for estimating the models, rather the five industry portfolios, we consider the robustness of our results when estimating the models at the quarterly frequency. Quarterly data is only available from 1947Q1 and our sample period becomes 1947Q1-2017Q1. In this case, the summary statistics for our data are altered, as shown in Table 13 (note that none of the figures presented in this subsection are annualised). In particular, the mean log market return is slightly higher, at around 1.8% per quarter (or 7.2% a year).

In addition, we must change the definitions of $\mathrm{Var}_t(r_{m,t+1})$ and $\bar{r}_{m,t}$ to reflect the fact we are using quarterly data. These become

$$\mathrm{Var}_t(r_{m,t+1}) = 3\sum_{d=1}^{12} w_d (r^m_{m,t-d} - \bar{r}_{m,t})^2; \qquad \bar{r}_{m,t} = \frac{1}{3}\sum_{d=1}^{3} r^m_{m,t-d}.$$

The definition of w_d remains the same and we continue to use 12 months of data to compute the conditional variance.

We estimate the models using both the optimal and identity weight matrices. Summary statistics for the residuals are shown in Table 14. Note that these are quarterly figures (one could annualise them by multiplying them by four). As we can see in Table 14, only the Cecchetti-Lam-Mark model estimated with the identity matrix provides a credible residual series and therefore a credible expected return series, with a mean residual of -0.8% per quarter. The Bansal-Yaron model certainly does not provide credible

Table 12: Campbell-Cochrane model results - Industry portfolios and identity matrix

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.142	-0.220	0.028	-0.040	0.039	-0.149	0.179	0.303	0.318
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.009	0.006	0.782	0.733	0.768	0.310	0.261	0.076	0.080

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.009	-0.006	-0.003	0.001	0.006	0.010
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.05	-0.014	-0.028	0.078	0.138	0.176	0.224	0.259	0.285	0.305
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.1	0.017	0.011	0.083	0.135	0.155	0.194	0.246	0.278	0.311
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.25	-0.022	-0.037	-0.027	-0.058	-0.082	-0.069	-0.065	-0.057	-0.064
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.5	-0.065	-0.111	-0.113	-0.158	-0.195	-0.210	-0.207	-0.208	-0.212
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.75	-0.029	-0.090	-0.121	-0.151	-0.154	-0.133	-0.113	-0.115	-0.116
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	-0.034	-0.047	-0.083	-0.099	-0.130	-0.148	-0.160	-0.183	-0.199
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	-0.030	-0.059	-0.086	-0.080	-0.083	-0.092	-0.106	-0.121	-0.139
	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.99	-0.009	-0.018	-0.026	-0.041	-0.059	-0.079	-0.099	-0.121	-0.144
	0.62	0.47	0.47	0.47	0.44	0.42	0.42	0.42	0.42

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.491	0.511	0.527	0.539	0.548	0.556	0.562	0.568	0.575
p-value	0.624	0.609	0.598	0.590	0.584	0.579	0.574	0.570	0.566

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.412	–0.022	182144
(Std err)	(0.528)	(0.035)	
p-value	0.439	0.522	0.000

(e) Rescaled range

$\widehat{Q}$	0.946
p-value	0.00

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one. In Panel (e), $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 13: Quarterly data summary statistics

	Mean	Median	Std dev	$\hat{\rho}(1)$
r_m	0.018	0.029	0.081	0.077
r_f	0.002	0.003	0.007	0.745
Δc	0.005	0.006	0.005	0.279
Δd	0.007	0.001	0.148	0.584
z_m	4.871	4.851	0.426	0.937

Descriptive statistics for our key variables at the quarterly frequency over the period 1947Q1-2017Q1. r_m denotes the log market return, r_f the quarterly log risk-free rate (the rolled over 1 month US T-bill), Δc log consumption growth, Δd log dividend growth and z_m the log price-dividend ratio. "Std dev" denotes standard deviation and "$\hat{\rho}(1)$" first-order serial correlation.

Table 14: Properties of $\hat{\xi}_t$ - Quarterly

Model	Mean	Median	Std dev	$\hat{\rho}(1)$
Optimal weight matrix				
Bansal-Yaron	-	-	-	-
Campbell-Cochrane	-0.351	-0.341	0.091	0.264
Cecchetti-Lam-Mark	-0.389	-0.886	0.726	0.871
Identity weight matrix				
Bansal-Yaron	-9.166	-9.172	0.231	0.791
Campbell-Cochrane	-0.203	-0.186	0.094	0.318
Cecchetti-Lam-Mark	-0.008	0.003	0.081	0.074

Summary statistics for the model-implied ex-ante residuals. "Std dev" denotes standard deviation and "$\hat{\rho}(1)$" first-order serial correlation. The models are estimated and residuals computed using quarterly US data over the period 1947Q1-2017Q1.

residual series: it has mean quarterly residuals of -9200% per quarter with the identity weight matrix and the GMM estimation does not converge with the optimal weight matrix. The Campbell-Cochrane model generates mean residuals of -35% per quarter with the optimal weight matrix and -20% per quarter with the identity weight matrix.

The results for the Cecchetti-Lam-Mark model estimated at the quarterly frequency with the identity weight matrix are in Table 15. Note that the Cecchetti-Lam-Mark state variable is affected by the change of data frequency. Note also that q indicates the horizon in quarters. The choice of $q = 8, 12, 16, 20, 24, 28, 32, 36, 40$ quarters aligns with the earlier choice of $q = 2, 3, 4, 5, 6, 7, 8, 9, 10$ years. There are no rejections of the MDS null for the residuals. Only the Mincer-Zarnowitz F-test rejects the model. Overall, these results seem to suggest that the Cecchetti-Lam-Mark model can explain the dynamics of returns. In addition, the maximal predictability results only show one significant exceedence of the R^2 bound. Note, however, that the R^2 bound exceeds one on three occasions, which may be a symptom of numerical issues in computing the bounds. Moreover, the semi-parametric MIDAS test fails to reject in favour of the Cecchetti-Lam-Mark state variable.

The findings that the Cecchetti-Lam-Mark model and its state variable may be able to explain return dynamics, however, are not themselves robust. Having a larger sample allows us to look at performance in sub-samples. We divide our sample in two with the break in the middle of the sample, so that our sub-samples are 1947Q1-1982Q1 and 1982Q2-2017Q1.

In addition, we can examine robustness to dealing with look-ahead bias in the second sub-sample. In the above results, the parameters of the ex-ante $(t-1)$ expectations are estimated over future data, which could induce a finite-sample bias in the test statistics even when the test statistics are asymptotically valid. These concerns apply only to the correlogram and Hong-Lee tests. The quantilogram and rescaled range bootstrap procedures explicitly account for the estimation method and the finite sample. The Mincer-Zarnowitz, MIDAS and maximum predictability tests condition on the parameter estimates in any case. We evaluate the robustness of our correlogram and Hong-Lee results to using past data only to estimate the parameters of the model residuals. We compute residuals for the second sub-sample which are formed using parameters estimated over an expanding window. The expanding window begins at the first observation in the whole sample (1947Q1) and ends at the $(t-1)$th observation when computing the $t-1$ expectations of returns at t. 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.

Looking at the Cecchetti-Lam-Mark residuals estimated with the identity matrix in the sub-samples in this way, we see that the MDS null is rejected in both sub-samples and when we account for look-ahead bias. The MDS null is clearly rejected by the quantilograms in the first sub-sample (Table 16a): 37 of the 81 weighted quantilograms are significant at the 10% level and 25 of those are significant at the 5% level. Untabulated results show that this is the only test to reject the null in the first sub-sample, reiterating why it is important to consider a battery of test statistics. Looking at the second sub-sample (Table 16b), the MDS null is easily rejected by the Hong-Lee tests. When accounting for look-ahead bias in the estimation (Table 16c), the MDS null remains strongly rejected, this time by the weighted correlograms. That the Mincer-Zarnowitz F-test provides the only rejection of the Cecchetti-Lam-Mark model in Table 15 is consistent with the rejections of the model in the second sub-sample since the Mincer-Zarnowitz regression is based on the out-of-sample expected returns for the second sub-sample only.

Table 15: Cecchetti-Lam-Mark model - quarterly results with identity weight matrix

(a) Correlogram

q	8	12	16	20	24	28	32	36	40
$\bar{C}(q)$	-0.029	-0.095	-0.068	-0.007	-0.017	-0.073	-0.038	-0.040	-0.055
(Std err)	(0.089)	(0.112)	(0.132)	(0.149)	(0.164)	(0.178)	(0.191)	(0.203)	(0.215)
p-value	0.739	0.396	0.607	0.965	0.916	0.682	0.840	0.843	0.796

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	0.007	0.003	-0.003	-0.008	-0.014	-0.020	-0.025	-0.030	-0.036
	0.57	0.78	0.85	0.95	0.78	0.66	0.48	0.39	0.35
0.05	0.044	0.054	0.054	0.049	0.038	0.024	0.013	0.006	0.000
	0.37	0.37	0.35	0.37	0.38	0.33	0.30	0.29	0.27
0.1	0.054	0.061	0.059	0.054	0.041	0.023	0.010	0.001	-0.006
	0.46	0.40	0.33	0.33	0.36	0.36	0.35	0.35	0.32
0.25	0.043	0.048	0.045	0.043	0.035	0.018	0.003	-0.006	-0.014
	0.42	0.41	0.41	0.39	0.36	0.32	0.34	0.36	0.33
0.5	0.008	0.009	0.005	0.001	-0.004	-0.008	-0.014	-0.020	-0.025
	0.62	0.56	0.59	0.60	0.51	0.48	0.45	0.41	0.38
0.75	0.033	0.034	0.032	0.026	0.014	-0.003	-0.019	-0.029	-0.036
	0.49	0.46	0.41	0.39	0.41	0.43	0.40	0.37	0.30
0.9	0.042	0.041	0.034	0.024	0.009	-0.011	-0.028	-0.041	-0.052
	0.43	0.45	0.44	0.46	0.53	0.53	0.55	0.49	0.43
0.95	0.045	0.049	0.044	0.034	0.020	0.004	-0.009	-0.018	-0.027
	0.43	0.36	0.38	0.37	0.37	0.39	0.41	0.37	0.33
0.99	0.007	0.007	0.004	0.000	-0.008	-0.018	-0.025	-0.030	-0.035
	0.77	0.99	0.90	0.73	0.66	0.62	0.53	0.46	0.39

(c) Hong-Lee tests

q	8	12	16	20	24	28	32	36	40
$\widehat{G}(q)$	0.002	-0.211	-0.367	-0.468	-0.542	-0.603	-0.682	-0.731	-0.769
p-value	0.998	0.833	0.714	0.640	0.588	0.547	0.495	0.465	0.442

(d) Mincer-Zarnowitz regression

	Constant	Expected return	MZ F-stat
Coefficient	0.004	0.055	1082
(Std err)	(0.018)	(0.049)	
p-value	0.837	0.269	0.000

$\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (d) shows the Mincer-Zarnowitz regression estimates, with the relevant HAC standard errors (in parentheses) and p-values below, along with the HAC F-statistic testing the joint null that the constant is zero and the expected return coefficient is one.

Table 15: Cecchetti-Lam-Mark model - quarterly results with identity weight matrix

(e) Rescaled range

$\widehat{Q}$	1.076
p-value	0.87

(f) SP MIDAS test

$\widehat{\mathcal{J}}$	-0.725
p-value	0.602

(g) 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	5.6×10^{-7}	0.499	3.071	3.196	0.861
Wald stat	-	-	-	-	10.95	-	-	-	-
p-value	-	-	-	-	0.001	-	-	-	-

$\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath. $\widehat{\mathcal{J}}$ is the estimated Hsiao et al. (2007) consistent model specification statistic from the semi-parametric MIDAS model. Its bootstrapped p-value is given beneath. Panel (e) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$). The Wald statistic and its asymptotic p-value are reported.

Moreover, there are three significant exceedences of the R^2 bound in each sub-sample, although not necessarily at the same horizons. The R^2 bound is significantly exceeded at $q = 28$ in both sub-samples, but not the whole sample. The ability of the Cecchetti-Lam-Mark model state variable to explain the dynamics of returns also appears not to be robust. In addition, the semi-parametric MIDAS test fails to reject in favour of the Cecchetti-Lam-Mark state variable in either sub-sample (sub-sample 1: $\widehat{\mathcal{J}} = -0.721$, p-value $= 0.549$; sub-sample 2: $\widehat{\mathcal{J}} = -0.719$, p-value $= 0.566$).

We lastly consider the robustness of the Bansal-Yaron and Campbell-Cochrane state variable test results to using quarterly data. Note that the Bansal-Yaron state variables do not depend on whether we estimate the Bansal-Yaron model using the identity or optimal weight matrix, but the Campbell-Cochrane state variables do depend on the weight matrix used.

Table 17 shows the results of the semi-parametric MIDAS test robustness checks, while Table 18 shows the results of the maximal predictability robustness checks. For the Bansal-Yaron model we see a rejection of the null that the state variables are not relevant for expected returns conditioning on $\mathrm{Var}_t(r_{m,t+1})$ in the semi-parametric MIDAS test. However, the maximal predictability tests give similar results to when using the annual data, suggesting that the model's state variables cannot explain the own-history predictability of returns. So the picture is mixed.

For the Campbell-Cochrane model, things look a little more hopeful. The semi-parametric MIDAS test clearly rejects in favour of the Campbell-Cochrane state variable. However, there are two significant exceedences of the R^2 bound using the optimal weight matrix and three using the identity weight matrix at the lag lengths considered. Moreover, untabulated results show a number further rejections at horizons $q < 8$ in both cases. Using the optimal weight matrix, the R^2 bound is exceeded for $q = 3, 4, 5$ and 6 quarters and these exceedences are significant at the 1% level. Using the identity weight matrix, there are significant exceedences for $q = 1$ and 6 quarters. Again, the quarterly data give a mixed picture on the Campbell-Cochrane state variable.

We take these maximal predictability results with a little caution, however. Table

Table 16: Cecchetti-Lam-Mark model quarterly sub-sample results using identity weight matrix

(a) Quantilogram - sub-sample 1: 1947Q1-1982Q1

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	0.050	0.076	0.100	0.122	0.142	0.163	0.193	0.337	0.589
	0.02	0.02	0.02	0.06	0.07	0.09	0.09	0.65	0.62
0.05	0.092	0.093	0.075	0.076	0.065	0.042	0.010	-0.009	-0.036
	0.82	0.97	0.86	0.75	0.62	0.48	0.33	0.24	0.19
0.1	0.008	-0.002	-0.017	0.002	0.004	-0.006	-0.012	-0.007	-0.018
	0.52	0.41	0.40	0.49	0.45	0.42	0.35	0.35	0.29
0.25	-0.095	-0.147	-0.181	-0.189	-0.198	-0.212	-0.226	-0.239	-0.244
	0.06	0.04	0.03	0.03	0.03	0.03	0.02	0.02	0.01
0.5	-0.052	-0.133	-0.191	-0.207	-0.200	-0.203	-0.221	-0.236	-0.232
	0.30	0.10	0.07	0.05	0.04	0.04	0.03	0.02	0.02
0.75	0.036	-0.049	-0.111	-0.136	-0.161	-0.195	-0.227	-0.234	-0.225
	0.84	0.26	0.10	0.07	0.04	0.03	0.01	0.01	0.01
0.9	0.063	0.060	0.043	0.013	-0.039	-0.095	-0.140	-0.167	-0.186
	0.97	0.75	0.59	0.38	0.23	0.16	0.11	0.09	0.08
0.95	0.008	0.011	-0.002	-0.025	-0.066	-0.107	-0.142	-0.166	-0.183
	0.57	0.49	0.41	0.33	0.22	0.19	0.17	0.12	0.09
0.99	-0.029	-0.046	-0.059	-0.073	-0.091	-0.108	-0.128	-0.153	-0.184
	0.16	0.17	0.17	0.18	0.17	0.07	0.01	0.00	0.00

(b) Hong-Lee tests - sub-sample 2: 1982Q2-2017Q1

q	8	12	16	20	24	28	32	36	40
$\widehat{G}(q)$	20.29	19.83	19.52	19.30	19.08	18.78	18.46	18.15	17.87
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

(c) Correlogram - sub-sample 2: 1982Q2-2017Q1, accounting for look-ahead bias

q	8	12	16	20	24	28	32	36	40
$\bar{C}(q)$	-3.548	-11.02	-26.13	-35.71	196.3	-39.16	-38.37	-13.32	21.03
(Std err)	0.251	0.318	0.373	0.422	0.465	0.504	0.541	0.575	0.608
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Panel (a) reports the quantilogram tests of the MDS null for the Cecchetti-Lam-Mark model residuals estimated with the identity weight matrix, over the first sub-sample 1947Q1-1982Q1. The estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. Panel (b) gives the Hong-Lee tests for the residuals from the second sub-sample 1982Q2-2017Q1. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (c) reports the weighted correlogram tests for the second sub-sample where estimation uses the identity weight matrix but also accounts for possible look-ahead bias. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1 - j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath.

Table 16: Cecchetti-Lam-Mark model quarterly sub-sample results using identity weight matrix

(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	-	-	-	-	494.3	806.1	553.0	-	-
p-value	-	-	-	-	0.000	0.000	0.000	-	-
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	442.1	167.2	-	924.5	-	-	13078
p-value	-	0.380	0.000	0.000	-	0.000	-	-	0.000

Panel (d) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$) in each of the two sub-samples. The Wald statistic and its asymptotic p-value are reported.

Table 17: Quarterly semi-parametric MIDAS results

Model	$\widehat{\mathcal{J}}$	p-value
Bansal-Yaron	1.471	0.048
Campbell-Cochrane		
Optimal weight matrix	0.559	0.018
Identity weight matrix	0.851	0.018

$\widehat{\mathcal{J}}$ is the estimated Hsiao et al. (2007) consistent model specification statistic from the semi-parametric MIDAS regression for the asset pricing model and estimation method specified. Its bootstrapped p-value is given in the final column. The Bansal-Yaron model is estimated using the identity weight matrix. Both models are estimated using quarterly data over the period 1947Q1-2017Q1.

Table 18: Quarterly maximal predictability results

q	8	12	16	20	24	28	32	36	40
Bansal-Yaron model									
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	216041	148018	-	267587	416715	-	593143	1307207	-
p-value	0.000	0.000	-	0.000	0.000	-	0.000	0.000	-
Campbell-Cochrane model - optimal weight matrix									
R^2	0.029	0.041	0.059	1.2×10^{-4}	0.001	0.005	1.4×10^{-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	462.1	7.954	-	-	-	-	-	-	-
p-value	0.000	0.005	-	-	-	-	-	-	-
Campbell-Cochrane model - identity weight matrix									
R^2	-30.42	0.022	0.034	0.445	0.082	0.085	1.5×10^{-6}	0.025	1.000
$\bar{R}^2$	1225	0.622	4.4×10^{-4}	0.300	0.903	7.0×10^{-5}	0.831	0.335	9.772
Wald stat	-	-	25.26	8.190	-	197.0	-	-	-
p-value	-	-	5.0×10^{-7}	0.004	-	0.000	-	-	-

Tests of the null that the market return is no more predictable than implied by the Bansal-Yaron/Campbell-Cochrane model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1947Q1-2017Q1. The Wald statistic and its asymptotic p-value are reported.

18 shows that there are numerical difficulties in estimating the R^2 and $\bar{R}^2$ parameters. These are estimated jointly by GMM (no regression is run to obtain R^2). As a result, even though the R^2 for the predictive regressions should be the same for both models and whether the optimal or identity weight matrix is used to estimate the model, this is not the case. Moreover, we see some R^2 and $\bar{R}^2$ which are either greater than one or less than zero. These numerical issues may be a function of the mis-specification of the state variables in terms of being able to explain own-history predictability of returns. Or they may reflect more general numerical issues.

7 Conclusion

We show that three consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the dynamics of the US market return. First, we estimate the models and derive ex-ante expected returns from them. The difference between the expected returns and realised returns is not MDS, which it should be if the models are correctly specified, due to rational expectations. Second, (Mincer and Zarnowitz, 1969) regressions show that out-of-sample expected returns generated from the models are systematically biased. This, again, would be ruled out by rational expectations if the model in question were correctly specified. Third, we show that the degree of predictability in the market return is not consistent with the state variables of any of the three models we test.

The failure of the models considered to capture the dynamics of stock returns has several different interpretations. The first is that perhaps some auxiliary assumption in the models has failed. For example, the assumed joint normality of consumption and dividend growth in the Campbell-Cochrane model (used to derive expected returns) or the assumed joint normality of consumption growth, dividend growth, the long-run risk and economic volatility in the Bansal-Yaron model (used by Constantinides and Ghosh

(2011) to invert the model and derive the moment conditions to estimate it). Even if everything else were correctly specified in the model, the expected returns would still be compromised, as their extraction relies on the normality assumption. Note that these normality assumptions are used when backing out the state variables for the maximal predictability and MIDAS tests too, so the state variable tests would be affected in this scenario as well. In this interpretation, the models are basically correct, but the auxiliary assumptions need to be relaxed to get expected returns consistent with the dynamics of the market return.

A second interpretation in which the models are basically correct is to say that the models presented are equilibrium models, but that financial markets are often out of equilibrium. Therefore, to model market dynamics, it is necessary to consider a framework in which markets adjust to a (possibly time-varying) equilibrium. Adam et al. (2016) present such a model. They have an agent with CRRA preferences who knows the risk-adjusted stock 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 serial correlation and excess volatility upon the returns, even when 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 account for the equity premium and risk-free rate puzzles. Nonetheless, it is possible that by applying this framework to, for example, the Campbell-Cochrane model would account for these puzzles.

Finally, it may simply be that the state variables the models use to decribe the SDF, and therefore risk, are mis-specified. More state variables may need to be considered, some of those considered may need to be dropped, or some combination of the two. In this interpretation, the properties of aggregate stock returns can still be explained by risk in a rational framework. The issue is that we do not have a suitable definition of risk. A further interpretation would be that the properties of stock returns cannot be explained by risk alone, and that some kind of systematic irrationality or bias on the part of investors, along with the barriers to arbitrage that prevent rational investors from eliminating mis-pricing, need to be considered.

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

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

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

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

In the stochastic discount factor

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

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 κ_0 and κ_1 derive from applying the Campbell and Shiller (1988) log-linearisation procedure to the returns to the consumption claim and market portfolio (Bansal and Yaron, 2004). 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/dividend 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:

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

A.3 Expected return coefficients

The expected market return in the Bansal-Yaron model is

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

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.

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`.

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. 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.

Cecchetti, S. G., Lam, P.-S., and Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review*, 80(3):398–418.

Cheng, X., Dou, W. W., and Liao, Z. (2022). Macro-finance decoupling: Robust evaluations of macro asset pricing models. *Econometrica*, 90(2):685–713.

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). Beyond arbitrage: good deal asset price bounds in incomplete markets. *Journal of Finance*, 108(1):79–119.

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.

Delgado, M. A. and Velasco, C. (2011). An asymptotically pivotal transform of the residuals sample autocorrelations with application to model checking. *Journal of the American Statistical Association*, 106(495):946–958.

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.

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`.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Hong, Y. (1999). Hypothesis tests in time series via the empirical characteristic function: a generalised spectral density approach. *Journal of the American Statistical Association*, 94(448).

Hong, Y. and Lee, Y.-J. (2005). Generalised spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form. *Review of Economic Studies*, 72(2).

Hsiao, C., Li, Q., and Racine, J. S. (2007). A consistent model specification test with mixed discrete and continuous data. *Journal of Econometrics*, 140(2):802–826.

Huang, D. and Zhou, G. (2017). Upper bounds on return predictability. *Journal of Financial and Quantitative Analysis*, 52(2):401–425.

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.

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. (1989). The size and power of the variance ratio test in finite samples: a Monte Carlo investigation. *Journal of Econometrics*, 40(2):203–238.

Lobato, I. N., Nankervis, J. C., and Savin, N. E. (2002). Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory*, 18(3):730–743.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1):255–285.

Mehra, R. and Prescott, E. C. (1985). The equity premium: a puzzle. *Journal of Monetary Economics*, 15(2):145 – 161.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887.

Mincer, J. A. and Zarnowitz, V. (1969). *The Evaluation of Economic Forecasts*, pages 3–46. NBER.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1):27–59.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49.

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.