NBER WORKING PAPER SERIES

MONETARY POLICY AND ASSET VALUATION:
EVIDENCE FROM A MARKOV-SWITCHING CAY

Francesco Bianchi
Martin Lettau
Sydney C. Ludvigson

Working Paper 22572
http://www.nber.org/papers/w22572

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

Ludvigson acknowledges research support from the C.V. Starr Center for Applied Economics at NYU. We thank Matteo Maggiori and Michael Weber for helpful comments. Any errors or omissions are the responsibility of the authors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Monetary Policy and Asset Valuation: Evidence From a Markov-Switching cay
Francesco Bianchi, Martin Lettau, and Sydney C. Ludvigson
NBER Working Paper No. 22572
August 2016, Revised January 2017
JEL No. E02,E4,E52,G12

ABSTRACT

This paper presents evidence of infrequent shifts, or “breaks,” in the mean of the consumption-wealth variable cay_t that are strongly associated with fluctuations in the long-run expected value of the Federal Reserve’s primary policy rate, with low policy rates associated with high asset valuations, and vice versa. By contrast, there is no evidence that infrequent shifts to high asset valuations and low policy rates are associated with higher expected economic growth or lower economic uncertainty; indeed the opposite is true. Additional evidence supports a “reaching for yield” channel, wherein low interest rate regimes coincide with low equity market risk premia.

Francesco Bianchi
Social Sciences Building, 201B
Department of Economics
Duke University
Box 90097
Durham, NC 27708-0097
and CEPR
and also NBER
francesco.bianchi@duke.edu

Martin Lettau
Haas School of Business
University of California, Berkeley
545 Student Services Bldg. #1900
Berkeley, CA 94720-1900
and NBER
lettau@haas.berkeley.edu

Sydney C. Ludvigson
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10002
and NBER
sydney.ludvigson@nyu.edu

1 Introduction

Asset values fluctuate widely around measures of economic fundamentals. Some of this variation is concentrated at high frequencies, as is apparent from the daily volatility in the stock market, and we might reasonably attribute the origins of this variation to market "noise" around a more stable aggregate economic state. But what if a significant fraction of this variation is attributable to low frequency, decades-long shifts in these relative relationships? Such a phenomenon, were it to exist, could not be readily attributed to short-term volatility in the stock market, but would instead raise a question about the role of structural changes in the macroeconomy that govern how high or low assets values can remain relative to economic fundamentals for prolonged periods.

This paper presents empirical evidence of just such a low frequency phenomenon. The consumption-wealth variable cay_t is an estimated asset market valuation ratio that uses data on total household net worth (asset wealth), but its variation is driven primarily by movements in the stock market relative to two key macroeconomic fundamentals: consumer spending and labor income. In a 2001 published paper, Lettau and Ludvigson (2001) (LL hereafter) introduced this variable and found that it had strong forecasting power for U.S. stock returns. We refer to this asset valuation variable interchangeably as a "wealth ratio," since it measures how high or low asset values are relative to a linear combination of consumption and labor income.

In the years since the 2001 paper was published, the statistical properties of the estimated cay_t series have shifted in some fundamental ways. Notably, the measured value of cay_t has become more persistent over time, resulting in forecasting power for stock market returns increasingly concentrated at longer horizons and making it difficult, according to some statistical tests, to distinguish cay_t from a unit root process. Mechanically, the reason for this has to do with the persistently high asset valuations of the post-millennial period, which have resulted in observations on cay_t that have remained well below the variable's pre-2000 mean even in the aftermath of two large stock market crashes and one large housing crash. Similar findings have been documented for other stock market valuation ratios long used as predictor variables for stock returns, including price-dividend or price-earnings ratios.[1] Despite these findings, a literal unit root interpretation for these variables is unappealing because it implies that stock prices or asset values could wander arbitrarily far from measures of fundamental value *indefinitely*.[2]

[1] The near-unit root statistical properties of these ratios and their implications for return forecasting have been the subject of empirical work by Lewellen (2004), Campbell and Thompson (2008), Lettau, Ludvigson, and Wachter (2008), Lettau and Van Nieuwerburgh (2008), van Binsbergen and Koijen (2010), and Koijen and Van Nieuwerburgh (2011).

[2] This is unappealing even with presumed departures from conventional notions of market efficiency. Even theories that postulate "bubbles" almost always imply that the bubble will eventually burst, restoring a pre-bubble relationship between prices and fundamentals.

An arguably more appealing interpretation is that there are instead infrequent shifts in certain moments of the stationary distribution that–when not taken into account–make distinguishing a stationary from a unit root variable difficult in a small sample.

This paper presents evidence that the highly persistent changes in asset values relative to fundamentals as measured by cay_t, including the persistently high asset valuations of the post-2000 period, are largely driven by infrequent shifts, or "breaks," in the mean of cay_t. In addition, these shifts are found to coincide with quantitatively large structural changes in the long-run expected value of the U.S. Central Bank's core policy interest rate, and with evidence for changes in risk-taking behavior in equity markets.

To establish this evidence, we first adjust cay_t for the infrequent shifts in its mean by estimating a Markov-switching version of the variable, denoted cay_t^{MS}. We find that the sample is divided into three clear subperiods characterized by two regimes for the mean of cay_t: a high mean/low asset valuation regime that prevails from 1976:Q2 to 2001:Q2, and a low mean/high asset valuation regime that prevails in two subperiods at the beginning and end of our sample, namely 1952:Q1-1976:Q1, and the post-millennial period 2001:Q2-2013:Q3. Unlike the conventional cay_t, which presumes a constant mean, cay_t^{MS} does not exhibit increasing persistence as estimates are updated over the sample to include the post-millennial period (the original estimation used data through 1998). Moreover, evidence in favor of stationarity for cay_t^{MS} is much stronger in current samples than it is for cay_t. This implies that infrequent shifts in the mean of cay_t largely explain why its statistical properties have shifted over time.

After documenting these findings, we first turn our attention to forecasts of the U.S. stock market. We find that the forecasting power of cay_t^{MS} for future stock market returns is superior to that of cay_t, even if no forward-looking data are used in the construction of cay_t^{MS}. This remains true out-of-sample, at least for some forecast horizons. Forecasts are improved because adjustments in the mean imply that estimates of conditional expectations do not mix data across regimes characterized by very different structural relationships between the level of cay_t and future asset returns.

We then direct our attention to the key question of what these infrequent mean shifts represent economically. Any estimated statistical relationship is subject to possible structural change as the number of years over which the relationship is measured rises. This may be especially true of cay_t, where the definitions of the embedded variables have changed discretely over time as data collection agencies have altered their measurement criteria for all three series. But structural shifts in the economy are also likely to play a role, as suggested by evidence that other stock market valuation measures have also experienced "breaks" in the mean values of their distributions (e.g., Lettau, Ludvigson, and Wachter (2008); Lettau and Van Nieuwerburgh (2008)). We are interested in the macroeconomic origins of these low frequency shifts in the mean of cay. Thus we estimate a macroeconomic vector autoregression for output growth, inflation, and the federal funds rate allowing for Markov-switching in its parameters (MS-VAR). Specifically, we allow the parameters of the MS-VAR to potentially undergo structural breaks during the periods that correspond to the shifts identified from our estimates of cay^{MS}. With this approach, we impose the formerly estimated regime sequence for cay on the macro MS-VAR, but the parameters characterizing the different regimes, as well as the transition matrix, are freely estimated. We find strong evidence of quantitatively large breaks in the long-run expected real federal funds rate that coincide with the breaks in the mean of cay, with low wealth ratios (low asset valuations or high cay) associated with an expectation of persistently high values for the real federal funds rate, and high wealth ratios (high asset valuations or low cay) associated an expectation of persistently low values for the real federal funds rate. The post-millennial period in particular, one characterized by relatively high asset valuations, is marked by expectations of prolonged low values for the real federal funds rate as compared to the middle subperiod where asset valuations were low and expected policy rates were high. By contrast, there is no evidence that these low frequency shifts to high asset valuations and persistently low policy rates are associated with higher expected economic growth over any horizon, or lower economic uncertainty; indeed the opposite is true. The findings therefore run counter to the idea that high asset valuations associated with a persistently low interest rate environment are the result of a positive outlook for economic growth, or lower uncertainty about that growth.

This latter result is perhaps not as surprising as it may first seem, if in fact monetary policy is largely responsible for the low frequency shifts in cay. After all, few macroeconomic theories imply that the Central Bank's ability to influence real activity stretches across the very long horizons over which the regime shifts we document occur. But it does raise a question as to why the Central Bank's expected policy rate is associated with such pronounced low frequency shifts in asset valuations. One answer is that the findings merely reflect corresponding regime shifts in discount rates, and indeed our evidence is consistent with this interpretation.

But even if we restrict attention to interpretations based on changing discount rates, theories differ on the reasons discount rates change with interest rates. Some theories tie low and declining discount rates entirely to the behavior of the risk-free rate, which is presumed to be driven endogenously downward by shocks that increase the fraction of wealth held in the hands of more risk averse or more pessimistic investors (e.g., Barro and Mollerus (2014); Caballero and Farhi (2014); Hall (2016)). In these theories, risk premia rise as the risk-free rate declines, implying that asset valuations can only be higher if the decline in the risk-free rate exceeds the rise in risk premia. Other theories imply that shifts downward in the risk-free rate coincide with shifts downward in risk-premia, as in models with a "reaching for yield" channel (e.g., Rajan (2006); Rajan (2013); Drechsler, Savov, and Schnabl (2014); Acharya and Naqvi (2016)).

Our empirical findings imply that the decline in discount rates in low interest rate environments is unlikely to be solely attributable to a downward shift in the expected risk-free rate large enough to more than offset an opposite movement in risk premia. Instead, we find evidence from equity markets that low interest rate regimes coincide with lower rather than higher risk premia, consistent with reaching for yield. Specifically, we find that, in a switch from a high to low interest rate regime, the estimated risk premia of 14 out of 17 portfolios that carry a positive risk premium on average simultaneously fall to lower levels. Moreover, both the risk premia and the book-market ratios (adjusted for expected earnings) of evidently more risky, higher Sharpe ratio portfolios, such as those that go long in value stocks or stocks that have recently appreciated the most, fall much more than those of evidently less risky, lower Sharpe ratio portfolios, such as those that go long in growth stocks or stocks that have recently appreciated the least. For several portfolios, the estimated risk premia reach lows or near-lows early in the post-2000 period and, after a brief spike upward in the 2007-08 financial crisis, again in the post-2009 period when interest rates entered the zero-lower-bound range. Our findings for stock market returns in this regard are reminiscent of recent evidence of reaching for yield in the Treasury market (e.g., Hanson and Stein (2015)), by U.S. prime money funds (e.g., Di Maggio and Kacperczyk (2015)), and by U.S. corporate bond mutual funds (Choi and Kronlund (2015)).

A growing empirical literature documents a linkage between monetary policymaking activities and financial returns in high frequency data, using either formal event studies and daily data (Cook (1989); Bernanke and Kuttner (2005); Gürkaynak, Sack, and Swanson (2005)) or by studying the timing of when premia in the aggregate stock market are earned in weeks related to FOMC-cycle time (Lucca and Moench (2015); Cieslak, Morse, and Vissing-Jorgensen (2015)) or by directly forecasting weekly stock returns using weekly observations on federal funds futures implied rates (Neuhierl and Weber (2016)). To the best of our knowledge, the findings presented in this paper are the first formal statistical evidence that low frequency, decades-long structural breaks in equity market return premia and asset values relative to economic fundamentals are strongly associated with long-horizon expectations of the primary policy instrument under direct control of the central monetary authority.

We emphasize that our results are silent on the question of whether the shifts expected policy rates we identify represent exogenous fluctuations, independent of aggregate economic conditions, or endogenous responses to these conditions. For example, if secular stagnation leads both agents and the monetary authority to expect a prolonged period of low growth, the Central Bank might be expected to accommodate this state by keeping policy rates accordingly low for an extended time. Whether or not the Central Bank's behavior is a truly exogenous impulse or an endogenous accommodation, our results imply that low frequency shifts in the stance of monetary policy have long-term implications for asset valuations.

The rest of the paper is organized as follows. The next section discusses the empirical model and the estimation of a Markov-switching cay_t. Section 3 presents results from this estimation, including evidence of breaks in the mean of cay, evidence on the persistence of cay once corrected for regime shifts in its mean, and a comparison of the forecasting power of cay^{MS} and cay for U.S. stock market returns. These investigations are all designed to address the question of whether or not infrequent shifts in the mean of cay_t can econometrically account for its changing statistical properties over the last 15 years. This is an important first step in understanding the macroeconomic sources of highly persistent fluctuations in asset market valuations, since if we cannot pinpoint the statistical mechanisms behind the changing behavior, we cannot hope to explain it. With this evidence in hand, section 4 turns to the key economic question of what macroeconomic forces lie behind the regime shifts in the mean of cay_t. Section 5 adds to this evidence by empirically evaluating the hypothesis that some of the shift to high asset valuation regimes is attributable to a decline in risk premia in equity market assets. Section 6 concludes.

2 Econometric Model of cay^{MS}

The variable studied by LL, denoted cay_t, is a stationary linear combination of log consumer spending, c_t, log asset wealth, a_t, and log labor income, y_t, all measured on an aggregate basis. Under assumptions described in LL and elaborated on in Lettau and Ludvigson (2010), cay_t may be interpreted as a proxy for the log consumption- aggregate (human and non-human) wealth ratio, and its relationship with future growth rates of a_t (highly correlated with stock market returns in quarterly data) and/or future growth rates of c_t and y_t, may be motivated from an aggregate household budget constraint. Theory implies that c_t, a_t, and y_t should be cointegrated, or that cay_t should be covariance stationary.

This section presents the econometric model of regime switches in the mean of cay_t. In the standard estimation without regime shifts in any parameters, the stationary linear combination of c_t, a_t, and y_t can be written

$$c_t = \alpha + \gamma_a a_t + \gamma_y y_t + \epsilon_t, \tag{1}$$

where the parameters to be estimated are α, γ_a, and γ_y. The residual ϵ_t is the stationary linear combination of these data, referred to as the cointegrating residual. In this paper, we estimate a Markov-switching version of this cointegrating relationship, analogously written as

$$c_t = \alpha_{\xi_t^\alpha} + \beta_a a_t + \beta_y y_t + e_t,$$

where the notation $\alpha_{\xi_t^\alpha}$ indicates that the value of the constant depends on the existence of a latent state variable, ξ_t^α, presumed to follow a two-state Markov-switching process with transition matrix $\boldsymbol{H}^\alpha$. This implies that the constant term $\alpha_{\xi_t^\alpha}$ can assume one of two discrete values, α_1 or α_2. The parameters β_a and β_y are the slope coefficients in this cointegrating relationship, analogous to γ_a and γ_y in the fixed coefficient regression (1). The residual e_t is again a stationary variable by assumption.

The standard approach to estimating a single cointegrating relation such as (1) is to run a dynamic least squares regression (DLS–Stock and Watson (1993)) that controls for leads and lags of the right-hand-side variables in order to adjust for the asymptotic inefficiencies attributable to regressor endogeneity. Let $\boldsymbol{z}_t$ be a 3×1 vector of data on c_t, a_t, and y_t, and k leads and k lags of Δa_t and Δy_t and let $\boldsymbol{Z}_t = (\boldsymbol{z}_t, \boldsymbol{z}_{t-1}, ..., \boldsymbol{z}_1)$ be a vector containing all observations obtained through date t. To estimate the parameters of this stationary linear combination we modify the standard fixed coefficient DLS regression to allow for shifts in the intercept $\alpha_{\xi_t^\alpha}$:

$$c_t = \alpha_{\xi_t^\alpha} + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i} \Delta a_{t+i} + \sum_{i=-k}^{k} b_{y,i} \Delta y_{t+i} + \sigma \varepsilon_t \tag{2}$$

where $\varepsilon_t \sim N(0,1)$. The parameters of the econometric model include the cointegrating parameters and additional slope coefficients $\beta = \left(\beta_a, \beta_y, b\right)'$, where $b = (b_{a,-k}, .., b_{a,k}, b_{y,-k}, .., b_{y,k})'$, the two intercept values α_1 and α_2, the standard deviation of the residual σ, and the transition probabilities contained in the matrix $\boldsymbol{H}^\alpha$. Collect these parameters into a vector $\boldsymbol{\theta} = \left(\beta, \alpha_{\xi_t^\alpha}, \sigma, H^\alpha\right)'$.

Absent regime changes, *cay* is defined as:

$$cay_t^{FC} = c_t - (\alpha + \gamma_a a_t + \gamma_y y_t) \tag{3}$$

where the superscript "FC" stands for "fixed coefficients" because the constant α is fixed over time. Notice that when we impose a single regime, the Markov-switching model collapses back to the specification originally used by LL. The variable cay_t^{FC} is the same as that defined in LL where it was denoted cay_t. For the purposes of his paper, we have added the superscript "FC" in order to explicitly distinguish it from the Markov-switching version. The parameters $\boldsymbol{\theta}$ of the time-series model for cay_t^{FC} include the cointegrating parameters γ_a and γ_y, the additional slope coefficients on the leads and lags in the DLS regression, and the single intercept value α.

Let T be the sample size used in the estimation accounting for leads and lags in the regression. For the Markov-switching model, the constant $\alpha_{\xi_t^\alpha}$ depends on the regime ξ_t^α. If the sequence $\xi^{\alpha,T} = \{\xi_1^\alpha, ..., \xi_T^\alpha\}$ of regimes in place at each point in time were observed, we could immediately compute cay_t^{MS}. Unfortunately, $\xi^{\alpha,T}$ is generally unobservable and needs to be inferred together with the other parameters of the model. It follows that the two values for the Markov-switching constant $\alpha_{\xi_t^\alpha}$ (α_1 and α_2) must be weighted by their probabilities at each point in time. For this purpose, we consider two estimates of the state probabilities distinguished as *filtered* or *smoothed* probabilities. Let $P\left(\xi_t^\alpha = i | \boldsymbol{Z}_t; \boldsymbol{\theta}\right) \equiv \pi_{t|t}^i$ denote the probability that $\xi_t^\alpha = i$

based on data obtained through date t and knowledge of the parameters $\boldsymbol{\theta}$. We refer to these as filtered probabilities. Smoothed probabilities reflect the information about the state at time t that can be extracted from the whole sample: $P\left(\xi_t^\alpha = i | \boldsymbol{Z}_T; \boldsymbol{\theta}\right) \equiv \pi_{t|T}^i$.[3] These measures of the regime probabilities may be used to construct two versions of a Markov-switching cay, based on using either smoothed or filtered probabilities. In both cases, the intercept coefficient for cay is a probability weighted average of the two intercepts, α_1 and α_2. As a benchmark, we use the smoothed probabilities for our baseline estimate and denote it cay_t^{MS}. When we use filtered probabilities, we use the notation cay_t^{MSfilt}. Thus, cay_t^{MS} is computed one of two ways:

$$cay_t^{MSfilt} = c_t - \left(\sum_{i=1}^2 \pi_{t|t}^i \alpha_i + \beta_a a_t + \beta_y y_t\right). \quad (4)$$

$$cay_t^{MS} = c_t - \left(\sum_{i=1}^2 \pi_{t|T}^i \alpha_i + \beta_a a_t + \beta_y y_t\right). \quad (5)$$

Note that the econometric model (2) permits regime switches only in the intercept parameter. (The Appendix discusses alternative specifications in which σ is also subject to regime switches.) This specification for cay_t^{MS} maintains the hypothesis that a stationary linear combination of c_t, a_t, and y_t exists, just as in the standard cointegration specification and consistent with the aggregate budget constraint motivation given in LL. The only difference is that, here, some of the variation in the cointegrating residual is explicitly modeled via regime switches in its mean. To see this simply rewrite the above expressions for cay_t^{FC} and cay_t^{MS} to be inclusive of the intercept term, i.e.,

$$\begin{aligned} cay_t^{MS} + \sum_{i=1}^2 \pi_{t|T}^i \alpha_i &= c_t - \beta_a a_t - \beta_y y_t \\ cay_t^{FC} + \alpha &= c_t - \gamma_a a_t - \gamma_y y_t. \end{aligned}$$

These expressions show that the intercept term may be interpreted as the mean of a cointegrating residual on the right-hand-side in each case. Thinking of the residual as an asset valuation (or wealth) ratio, we may interpret the intercept as the mean of the asset valuation ratio itself. A high α_i corresponds to a low valuation ratio, since the residual $c_t - \beta_a a_t - \beta_y y_t$ is high whenever the value of wealth a_t is low relative to the implied linear combination of c_t and y_t with which it is cointegrated.

2.1 Estimation

We use Bayesian methods with flat priors to evaluate the regression parameters in (2). We first search for the posterior mode using a maximization algorithm. The posterior of the model and the corresponding regime probabilities $\pi_{t|t}^i$ and $\pi_{t|T}^i$ are obtained by computing the likelihood

[3] In using the DLS regression (2) to estimate cointegrating parameters, we lose 6 leads and 6 lags. The Appendix on computing cay^{MS} explains how we filter the data to obtain esitmates of the regime probabilities over the whole sample.

using the Hamilton filter (Hamilton (1994)), and combining it with priors. Since we use flat priors, the posterior coincides with the likelihood. Our estimate of cay_t^{MS} is based on the posterior mode of the parameter vector $\boldsymbol{\theta}$ and the corresponding regime probabilities. We further use a Gibbs sampling algorithm to do inference on the parameters and compute additional statistics of interest. In particular, uncertainty about the parameters, or any desired moments of the model parameters, can easily be characterized using the Gibbs sampling algorithm to compute the posterior distribution of model parameters. The full statement of the procedure and sampling algorithm is given in the Appendix.

3 Results: Breaks in the Mean of the Wealth Ratio

We estimate the Markov-switching cointegrating relation described by (2) over the sample 1952:Q1-2013:Q3 using six leads and lags. Table 1 reports the parameter estimates, while Figure 1 reports the probability of regime 1 for the Markov-switching intercept $\alpha_{\xi_t^\alpha}$ based on the posterior mode parameter estimates. The 90% credible sets are obtained making 2,000,000 draws from the posterior using the Gibbs sampling algorithm described above. One in every one thousand draws is retained. We check convergence using the methods suggested by Geweke (1992) and Raftery and Lewis (1992).[4]

The sample is divided into three clear subperiods characterized by two regimes for α. Regime 1 is a high α regime with the posterior mode point estimate equal to $\hat{\alpha}_1 = 0.9186$. The low α regime 2 posterior mode estimate is $\hat{\alpha}_2 = 0.8808$. A high α regime for *cay* corresponds to a low valuation ratio for the stock market, analogous to a low price-dividend ratio (Lettau and Ludvigson (2001)). Thus we shall refer to *high α regime 1* as the *low asset valuation regime*, and *low α regime 2* as the *high asset valuation regime*. Figure 1 shows that the low asset valuation regime prevails for a prolonged period of time starting from 1976:Q2 to 2001:Q2. The smoothed probability that $\alpha = \hat{\alpha}_1$ is unity during this period. By contrast, the pre-1976 and post-2001 subsamples are high asset valuation regimes, where the probability that $\alpha = \alpha_1$ is equal to 0. These correspond to the subperiods 1952:Q1-1976:Q1, and 2001:Q2-2013:Q3, respectively.

Table 1 provides estimates of the difference between the high and low α and its distribution. The difference is positive and statistically significant, as exemplified by the third row of Table 1, which shows that a 90% credible set only contains non-zero and positive values for this difference.[5] The two regimes turn out to be very persistent and this is reflected in the estimates

[4] For Raftery and Lewis (1992) we target 90% credible sets, with a 1% accuracy to be achieved with a 95% minimum probaility. We initialize the Gibbs sampling algorithm making a draw around the posterior mode. Sims and Zha (2006) point out that in Markov-switching models it is important to first find the posterior mode and then use it as a starting point for the MCMC algorithm due to the fact that the likelihood can have multiple peaks.

[5] The Gibbs sampling algorithm is used to generate a distribution for the difference between the two means in the same manner it is used to generate a distribution for any parameter. For each draw from the joint

for the diagonal elements of the transition matrix H^{α}, also reported in Table 1.

The mode values for the other cointegration parameters are $\beta_a = 0.26$ and $\beta_y = 0.62$. These values are comparable with those originally obtained by LL using a fixed coefficient regression ($\gamma_a = 0.31$ and $\gamma_y = 0.59$). By contrast, Table 2 reports the parameter estimates for the fixed coefficient cointegrating relation over the extended sample used in this paper, where $\gamma_a = 0.12$ and $\gamma_y = 0.78$. Therefore, in our current sample, the fixed coefficient parameter estimates differ substantially from those reported in 2001. Bearing in mind that deviations from the cointegrating relation are the result of persistent but transitory movements in a_t rather than c_t or y_t (Lettau and Ludvigson (2004), Lettau and Ludvigson (2013)), these results suggest that the fixed-coefficient estimates of cay_t attempted to "compensate" for increasingly persistent deviations in a_t from its cointegrating relation with c_t and y_t, by progressively reducing the weight on a_t and increasing the weight on y_t. The instability in these point estimates is largely eliminated by allowing for discrete shifts in the mean of *cay*.

To give a visual impression of these regimes over time, Figure 2 plots $cay_t^{MS} + \sum_{i=1}^{2} \pi_{t|T}^{i} \alpha_i$ over time, which is the estimated Markov-switching *cay* from (5) *inclusive* of the intercept. As explained above, we can think about this variable as a valuation ratio. When its mean is high, asset valuation is low. Also plotted as horizontal lines are the values $\hat{\alpha}_1$ and $\hat{\alpha}_2$ that arise in each regime over the sample. The figure shows that this valuation variable fluctuates around two distinct means in three separate periods of the sample, a low mean in the early part of the sample, a high mean in the middle, and a low mean again in the last part of the sample. Therefore, we refer to Regime 1 as the *high mean/low valuation regime*, whereas Regime 2 is the *low mean/high valuation regime*.

3.1 Persistence of cay^{MS} versus cay^{FC}

The most salient change in the statistical properties of the estimated cay_t^{FC} series in the time since it was originally introduced is that it has become more persistent. This can be traced directly to the persistently high asset valuations of the post-millennial period, which caused a_t to persistently deviate from any estimated common trend with c_t and y_t. Even after two stock market crashes, cay_t^{FC} remains well above its full sample mean. Figure 3 plots the fixed coefficient cay_t^{FC} and the Markov-switching cay_t^{MS} as defined in (3) and (5), respectively. (Unlike Figure 2, these values subtract the estimated α or probability-weighted α, respectively.) The two vertical bars mark the beginning and the end of the time span during which the high α regime was most likely to be in place, according to the smoothed probability estimates. As Figure 3 shows, cay_t^{FC} exhibits persistent deviations from zero, especially during the last subperiod 2001:Q2-2013:Q3, which coincides with the second appearance in our sample of the

distribution of the model parameters, we compute the difference and store it. We may then compute means and/or medians, and error bands, as for any other parameter of interest.

high asset valuation regime in which a_t has been persistently high relative to its estimated common trend with c_t and y_t. The estimated cay_t^{FC} also exhibits persistent deviations from zero during the period starting around 1980 and ending in the early 2000s, roughly coinciding with the occurrence of the low asset valuation regime. Most of this subperiod was included in the original estimation of cay_t^{FC}, so it has contributed less to the *growth* in the persistence in the series since that timer.

The key question then is whether adjusting for mean breaks in cay_t^{FC} leads to a series that largely corrects for the growth in the persistence of cay_t^{FC} over time. It is evident from Figure 3 that cay_t^{MS} is quite different from cay_t^{FC} in that it does not exhibit such persistent deviations from its demeaned value of zero. The reason is that the persistent deviations are instead captured by low-frequency regime changes in the constant of the cointegrating relation.

To formalize this visual impression, the first column of Table 3 reports the first-order autoregressive coefficient estimate for the two versions of cay. The estimated autocorrelation coefficient for cay_t^{FC} is 0.94. The estimated first-order autocorrelation coefficient for cay_t^{MS} is 0.81, which is close to the 0.79 estimated coefficient reported in LL. Allowing for low frequency mean shifts in the cointegrating relation largely restores the estimated persistence of cay to its original values.

Several other tests are employed to assess the degree of persistence in cay_t^{MS} as compared to cay_t^{FC}. First, we apply an augmented Dickey-Fuller t test to the estimated cointegrating residuals. We applied this test to the two versions of cay, and for different lagged values of Δcay in the Dickey-Fuller regression. The test statistics and corresponding critical values are reported in Table 3 . According to this test, the null hypothesis of no cointegration is rejected for the cay_t^{MS} in every case, whereas the opposite is true for cay_t^{FC}.

Second, we examine low frequency averages of cay to gauge its persistence. Figure 4 is based on weighted averages that summarize low-frequency variability in a series. Specifically, following Muller and Watson (2008) and Watson (2013), the figure plots the "cosine transformations" of each version of cay

$$f_j = \sum_{t=1}^{T} \cos\left(j(t-0.5)\pi T^{-1}\right) cay_t \quad \text{for } j = 1, ..., k.$$

As Muller and Watson (2008) show, the set of sample averages $\{f_j\}_{j=1}^{k}$, capture the variability in cay for periods greater than $2T/k$, where T is the sample size. Thus, with $T = 247$ quarters, the $k = 12$ points plotted in Figure 4 summarize the variability in cay for periods greater than $2 * 247/12 = 41.1667$ quarters, or approximately 10 years. Smaller values of j correspond to lower frequencies, so values of f_j plotted for small j (e.g., $j = 1, 2, 3$) give the variability in cay_t at low frequencies, while values of f_j plotted for higher j (e.g., $j = 10, 11, 12$) give the variability in cay_t at higher frequencies. A series that is integrated of order zero, $I(0)$, corresponding to covariance stationary, displays roughly the same variability (same value of f_j) at all frequencies j. By contrast, a series that is more persistent than $I(0)$ displays higher variability at low frequencies, resulting in higher values of f_j for low j than for high j. Figure 4 shows that the cosine transformation of cay_t^{MS} displays a pattern much more consistent with an $I(0)$ series than that of cay_t^{FC}, which shows a clear spike at $j = 3$, corresponding to a period of roughly 41 years.

As a third way to evaluate the persistence of in cay_t^{MS} versus cay_t^{FC}, we consider a parameterization from a fractionally integrated model. We therefore assume $(1-L)^d cay_t = u_t$, where L is the lag operator and u_t is an $I(0)$ process. If cay_t is $I(0)$, then $d = 0$. If cay_t has a unit root, then $d = 1$. Non-integer values of $d > 0$ are fractionally integrated series that are more persistent than $I(0)$ but less persistent than $I(1)$. Figure 5 shows the estimated log likelihoods for $(1-L)^d cay_t^{MS}$ and $(1-L)^d cay_t^{FC}$ as a function of d. For cay_t^{MS}, the likelihood peaks at $d = 0$, while for cay_t^{FC}, the likelihood rises with $d > 0$ and peaks near $d = 1.2$.[6]

Although the statistical tests just considered imply that cay_t^{FC} is sufficiently persistent that it can be difficult to distinguish from a unit root process, it does not follow that cay_t^{FC} actually has a unit root. Tests of the null of no cointegration are known to have low power against the cointegration alternative when deviations from the common trend are stationary but highly persistent. For this reason, Park (1990), Park (1992), Han and Ogaki (1997), and Ogaki and Park (1997) developed tests for the null of *cointegration*, rather than no cointegration, which they argue are more appropriate when theory suggests the variables should be cointegrated.[7] But even if cay_t^{FC} is simply highly persistent but ultimately stationary, the resulting low frequency deviations from a fixed mean raise issues for forecasting. We turn to these forecasting implications next.

3.2 Forecasts of Excess Stock Market Returns

The variable *cay* has been used as a stock market forecasting variable because most of its variation has been driven historically by transitory fluctuations in a around more stable values for c and y. Thus when a is high relative to c and y, that signals lower values for future asset returns, rather than higher values for c and/or y. It is therefore important to understand whether or not adjusting for the mean breaks in *cay* improves its forecasting power. If so, it suggests that these fundamental relationships between asset wealth, consumption and labor

[6]Please refer to Appendix 6.1 for details about the estimation of the fractionally integrated model.

[7]These tests, as they apply to *cay* specifically, are described in detail in the online appendix to (Lettau and Ludvigson (2013)) available on the authors' websites. Updated output from the Ogaki and Park (1997) test for the null of cointegration for cay_t^{FC} is provided in Table 4. As in earlier samples, this test continues to show no evidence against the null of cointegration for cay_t^{FC}, lending support to the hypothesis that the standard *cay* is stationary even if it is sufficiently persistent so as to make it difficult to distinguish from a non-stationary variable in our sample.

income exist within regimes but not across regimes. If not it suggests a breakdown in the relationship entirely.

Table 5 reports the results of long-horizon forecasts of log returns on the CRSP value-weighted stock market index in excess of a three month Treasury bill rate. This is the same return variable that was the focus of the empirical results in LL. The table compares the forecasting power of cay_t^{FC}, cay_t^{MSfilt}, based on filtered probabilities and cay_t^{MS}, based on smoothed probabilities. The top panel reports full sample forecasts. The bottom panel reports the results of forecasts based on fully recursive estimates of cay_t using data only up to time t. The recursive estimates are obtained as follows. First, all parameters $\boldsymbol{\theta}$ for each model are estimated in an initial period using data available from 1952:Q1 through 1980:Q4. All parameters are then reestimated recursively on data from 1952:Q1-1981:Q1, 1952:Q1-1981:Q2, and so on, until the final recursive estimate of cay is obtained based on data over the full sample 1952:Q1-2013:Q3. The recursively estimated values of cay_t^{FC}, cay_t^{MS}, are denoted cay^{FCrec} and cay^{MSrec}, respectively. These variables are then used to forecast returns over the entire subsample from 1981:Q1-2013:Q3. Notice that the recursive estimates use no forward looking data to estimate any of the parameters, including the regime probabilities, regimes values, or transition probabilities. In both panels we report the coefficient estimates on the regressor, the Newey and West (1987) corrected t-statistic, and the adjusted R^2 statistic.

The top panel shows that all measures of cay estimated over the full sample have statistically significant forecasting power for future excess stock market returns over horizons ranging from one to 16 quarters. But the coefficients, t-statistics and R^2 values are all larger using the Markov-switching versions cay_t^{MSfilt} and cay_t^{MS} than they are for cay_t^{FC}. The comparison is more stark if we compare recursively estimated values of cay to full sample values. For example, the full sample estimate of cay_t^{FC} explains 21% of the 16 quarter-ahead log excess stock market return in the subsample 1981:Q1-2013Q3, while cay_t^{MSrec} explains 42%. Moreover, in this subsample, cay_t^{FC} has little forecasting power for excess returns at all but the longest horizon, whereas cay_t^{MSrec} has much stronger forecasting power. These results show that accounting for infrequent shifts in the mean of cay_t delivers a more powerful predictor variable for returns, even if no forward looking data is used to form knowledge of the size and dates of the regime breaks.

Table 5 also shows that cay_t^{FCrec} also has stronger predictive power than cay_t^{FC} over this subsample. By recursively estimating the parameters in cay_t^{FC}, we allow them to change *every period*. In this way, a recursively estimated fixed-coefficient model can "compete" with the Markov-switching version, which explicitly models shifts in the mean parameter. The recursive estimation effectively allows the parameters of cay_t^{FCrec} (including the mean) to vary over different regimes of the sample. But finding that cay_t^{FCrec} performs better than cay_t^{FC} in forecasting returns hardly provides support for the hypothesis that the fixed-coefficient model is a better description of the data than the Markov-switching model. On the contrary, this finding may be taken as additional evidence of the instability in the fixed-coefficient parameters. If there were no such instability, cay_t^{FCrec} would be identically equal to cay_t^{FC}. Furthermore, because cay_t^{MS} is much less persistent than cay_t^{FC}, it is less subject to the spurious regression concerns raised by Ferson, Sarkissian, and Simin (2003) for return forecasts.

Table 6 reports mean-square forecast errors (MSEs) from out-of-sample forecasts. The forecasting relation is estimated in an initial period using data available from 1952:Q1 through 1980:Q4. Forecasts over the next h quarters are computed and forecast errors stored. The forecasting relation is then reestimated in rolling subsamples moving forward, (i.e., over the period 1952:Q1 through 1981:Q1), and forecasts and forecast errors are computed over the next h periods. This process is repeated until the end of the sample. Table 6 reports MSEs for several forecasting regressions. To form a basis for comparison, the first row reports results using nothing more than an estimated constant as a predictor variable, while the second row uses the lagged log excess stock market return as a predictor. The next three rows report results using cay_t^{FC}, cay_t^{MSfilt}, and cay_t^{MS} as univariate predictors. These versions of cay are all estimated using the full historical sample, as explained above. The last two rows report results using cay_t^{FCrec} and cay_t^{MSrec} as univariate predictors. These versions of cay are all estimated using only data up to and including date t, as explained above.

Some researchers have argued that many predictor variables for stock market returns have difficulty beating the sample mean of stock returns in out-of-sample tests (e.g., Goyal and Welch (2003); Goyal and Welch (2008)). The first row of Table 6 shows this is not the case here: all versions of cay have substantially lower out-of-sample MSEs than a forecasting model that uses only the (constant) sample mean of excess returns as a predictor, and even the recursively estimated versions have MSEs that are almost 70% smaller than those of the sample mean model. Table 6 shows that all versions of cay also have lower lower MSEs than a simple autoregressive forecasting model. Among those versions that are estimated using the full sample, the two Markov-switching versions, cay_t^{MSfilt}, and cay_t^{MS}, are much better predictors than the fixed-mean version cay_t^{FC}, having MSEs that are almost 50% smaller for 16-quarter return forecasts. The recursively estimated versions cay_t^{FCrec} and cay_t^{MSrec} have about the same predictive power over most horizons, although the Markov-switching cay offers a slight improvement over the fixed-mean cay at the longest (16 quarter) horizon. Because these recursive versions are estimated over short subsamples, the estimates of parameters are much noisier than they are for the full-sample versions, so it is not surprising that they have higher MSEs. For this very reason, it is likewise notable that cay_t^{MSrec} preforms as well (and slightly better at long horizons) as cay_t^{FCrec}, given that the former has many more parameters that require estimation over short subsamples of our quarterly dataset. Postwar samples of the size currently available are, however, much larger than the repeated subsamples used to construct the recursive estimates for this exercise. Going forward, such samples should provide less noisy estimates of cay_t^{MS} parameters. Researchers using cay_t as a predictor variable may wish to consider both measures as predictors of long-horizon stock market returns.[8]

4 What's Behind the Breaks in Asset Valuation?

For the rest of this paper, we search for empirical explanations for the breaks observed in *cay*. We argue that the low frequency nature of these shifts in wealth relative to consumer spending and labor earnings raise a question about the role of structural changes in the macroeconomy. We therefore carry out our investigation using macroeconomic data, studying how structural changes in the statistical properties of these data might be connected to the documented breaks in asset valuations as measured by *cay*.

To do so, we estimate a Markov-switching macroeconomic VAR (MS-VAR) for output growth, inflation, and the federal funds rate, allowing the parameters of the VAR to potentially undergo structural breaks during the periods that correspond to the shifts identified in our estimates for cay^{MS}. Specifically, we impose the formerly estimated regime sequence for *cay* on the VAR, but the parameters characterizing the different regimes, as well as the transition matrix, are freely estimated.[9] We denote the MS-VAR transition matrix $\boldsymbol{H}^A$ in order to distinguish it from the cay^{MS} transition matrix $\boldsymbol{H}^{\alpha}$. Note that the goal is not to estimate the regimes of the MS-VAR and see if they are aligned with the previously estimated breaks in *cay*. Instead, the goal is to establish what, if anything, is different in the VAR across the two previously estimated regimes that could help explain the breaks in the mean of *cay*. Thus we deliberately "tie our hands" by forcing the regime sequence for the MS-VAR to correspond to breaks in *cay*. We then ask whether the parameters of the MS-VAR show any evidence of important structural shifts under this sequence, when they are freely estimated and could in principle show no shift.

All MS-VARs estimated in this section and the next are implemented using Bayesian methods with flat priors. The Appendix provides estimation details.[10]

We consider the following MS-VAR model with n variables and $m = 2$ regimes:

$$Z_t = c_{\xi_t} + A_{1,\xi_t} Z_{t-1} + A_{2,\xi_t} Z_{t-2} + V_{\xi_t} \varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{6}$$

where Z_t is an $n \times 1$ vector of variables, c_{ξ_t} is an $n \times 1$ vector of constants, A_{l,ξ_t} for $l = 1, 2$ is an $n \times n$ matrix of coefficients, $V_{\xi_t} V'_{\xi_t}$ is an $n \times n$ covariance matrix for the $n \times 1$ vector of

[8]Both series are updated and available on the authors' websites.

[9]To impose the formerly estimated regime sequence, we choose the particular regime sequence $\hat{\xi}^{\alpha,T} = \{\hat{\xi}_1^{\alpha}, ..., \hat{\xi}_T^{\alpha}\}$ that is most likely to have occurred, given our estimated posterior mode parameter values for $\boldsymbol{\theta}$. See the Appendix for details.

[10]Bayesian methods are used because they offer significant computational advantages in characterizing uncertainty about parameter transformations such as risk-premia.

shocks ε_t. The process ξ_t controls the regime that is in place at time t and assumes two values, 1 and 2, based on the regime sequence identified in our estimates for cay^{MS}.

The vector Z_t includes three variables at quarterly frequency: GDP growth, Inflation, and the federal funds rate (FFR). Inflation and real output growth are defined as the year-to-year differences of the logarithm of the GDP price deflator and real GDP, respectively. The quarterly FFR is obtained by taking the average of monthly figures of the Effective Federal Funds Rate. All variables are taken from the FRED II database of the Federal Reserve Bank of St. Louis and are expressed in percentage points. The sample for this estimation spans the period 1955:Q3-2013:Q3. (The beginning of the sample is three years later than the sample used to estimate cay because the federal funds rate data is only available starting in 1955:Q3.) Details about the estimation can be found in the Appendix.

We are interested in knowing the conditional expectation and the conditional standard deviation of each variable in the MS-VAR as well as for the implied real interest rate RIR, defined as the difference between the FFR and one-step-ahead inflation expectations. For each variable $z_t \in Z_t$, the conditional expectation and conditional standard deviation are given by $\mathbb{E}_t\left(z_{t+s}\right)$ and $sd_t\left(z_{t+s}\right) = \sqrt{\mathbb{V}_t\left(z_{t+s}\right)} = \sqrt{\mathbb{E}_t\left[z_{t+s} - \mathbb{E}_t\left(z_{t+s}\right)\right]^2}$, where $\mathbb{E}_t\left(\cdot\right) \equiv \mathbb{E}\left(\cdot|\mathbb{I}_t\right)$ and $\mathbb{I}_t$ denotes the information available at time t. We assume that $\mathbb{I}_t$ includes knowledge of the regime in place at time t, the data up to time t, Z^t, and the VAR parameters for each regime. Both statistics are computed from the MS-VAR parameters and transition matrix $\boldsymbol{H}^A$, taking into account that future regimes are unknown and that there exists an entire posterior distribution of VAR parameters and transition matrix $\boldsymbol{H}^A$, translating into posterior distributions for $\mathbb{E}_t\left(z_{t+s}\right)$ and $sd_t\left(z_{t+s}\right)$. (Details on how these are calculated can be found in Appendix.) Note that inflation expectations are computed using the MS-VAR estimates, therefore RIR is not included directly in the MS-VAR but derived ex-post based on the MS-VAR estimates. Note that $sd_t\left(z_{t+s}\right)$ can be considered a measure of economic uncertainty, as implied by the MS-VAR.

Figure 6 reports the conditional expectations of each variable in the MS-VAR plus RIR. The figure reports the median and 68% credible sets from the posterior distribution of the conditional expectations. The figure shows striking evidence of structural change in the expected long-run RIR that coincide with the regime sequence estimated for the mean of cay. The occurrence of the low asset valuation regime in the middle subsample from 1976:Q2-2001:Q2, coincides with an expectation of sharply higher values for the real federal funds rate, while the periods of high asset valuation at the beginning (1955:Q3-1976:Q1) and end (2001:Q3-2013:Q3) of our sample coincide with expectations of much lower real interest rates. The differences across subsamples are strongly statistically significant: the figure also reports the 68% posterior credible sets for the conditional expectations and shows that the sets for the two regimes never overlap. Note that, because the MS-VAR parameters are freely estimated, the estimation could have found no evidence of structural change in the expected real interest rate across these subsamples and/or that changes occur in variables other than the expected real interest rate.

Figure 6 also shows that the estimated regime shifts in the expected future RIR show up most prominently in the expectations for the real policy rate five to ten years ahead. This finding underscores the extent to which low frequency shifts in the mean of *cay* coincide with expectations of a persistent low or high interest rate environment, rather than transitory movements in these rates.

There is no clear pattern with inflation. Thus the breaks in the expected real interest rate five to ten years ahead appear mostly attributable to breaks in the conditional expected value of the *nominal* interest rate, which the Federal Reserve directly influences. Of course the Federal Reserve may also have considerable influence over expected inflation. But movements in expected inflation do not line up as well with the regime sequence for breaks in the mean of *cay* as do movements in the expected nominal interest rate: in the first subperiod, corresponding to the first instance of the high asset valuation regime, expected inflation was low and then high, while in the second subperiod, corresponding to the low asset valuation regime, inflation was high and then low, where it remained throughout the entire span of the third subperiod corresponding to the second instance of the high asset valuation regime. To the best of our knowledge, these findings provide among the first formal statistical evidence that low frequency shifts in asset values relative to economic fundamentals are strongly associated with expectations of the long-run value of a policy instrument under direct control of the central monetary authority.

Why are high asset valuations associated with low expected long-run policy rates, and vice versa? Some classic theories of rational bubbles suggest that higher policy rates can lead to *higher* asset valuations (e.g., Galí (2014)). But the evidence here is inconsistent with this story, since high wealth ratios are associated with low policy rates rather than high. An alternative explanation, consistent with the evidence here, is that any change in the expected short-term real interest rate will always have some effect on asset values because it changes the "fundamental" value of the asset. If prices are sticky and the Federal Reserve reduces the nominal interest rate, changes in monetary policy may reduce the rate at which investor's discount future payouts by reducing the real "risk-free" rate component of the discount rate, thereby increasing asset values. This effect would also be present in bubbles of the resale-option variant, since unlike the classic rational bubble, the resale-option bubble is proportional to fundamental value (e.g., Harrison and Kreps (1978); Scheinkman and Xiong (2003)). Regardless of whether a bubble of this form is present or not, asset valuations would be further increased if the risk premium component of the discount rate falls simultaneously with the risk-free rate because investors' willingness to tolerate risk is for some reason inversely related to the long-run expected value of the Federal Reserve's core policy instrument, consistent with a "reaching for yield" channel. We present tests of this hypothesis in the next section.

Alternatively and/or in addition, high wealth ratios could be associated with low expected long-run policy rates because the latter are expected to generate either faster long-run economic growth, or lower uncertainty about that growth. Conversely, regimes characterized by low wealth ratios and high expected policy rates could be explained by lower expectations for long-run growth and/or higher uncertainty about that growth. Figure 6, however, provides no evidence that the low frequency shifts to high asset valuation regimes are associated with higher expected economic growth, or vice versa; indeed the opposite is true. The high asset valuation subperiods at the beginning and end of our sample are associated with *lower* expected GDP growth 10 years ahead than is the low asset valuation subperiod in the middle of the sample.

In theory, higher asset valuations could be the result of lower expected economic uncertainty (e.g., Lettau, Ludvigson, and Wachter (2008)). Figure 7 reports the conditional standard of each variable in the MS-VAR plus RIR. The figure reports the median and 68% credible sets from the posterior distribution of the conditional standard deviations. Note that the conditional standard deviation represents a statistical measure of uncertainty. The result in Figure 7 shows that for GDP growth and inflation, uncertainty is *higher* rather than lower in subperiods of the high asset valuation regime as compared to the subperiod of the low asset valuation regime, but the opposite is true for the nominal and real federal funds rate. Thus infrequent shifts to high mean wealth ratios cannot be readily explained by lower economic uncertainty, nor can it be explained by lower inflation uncertainty. Therefore, we can conclude that the high asset valuation regime is characterized by higher uncertainty for real activity and inflation, but lower uncertainty about the Federal Reserve's policy instrument, while the converse is true for the low asset valuation regime. The finding that macro and fed funds rate uncertainty vary inversely across the regimes is consistent with a more active role of the Federal Reserve in stabilizing inflation and real activity. As the Federal Reserve is expected to respond more aggressively by raising interest rates to counter higher inflation and/or a lower output gap, macroeconomic volatility is reduced, whereas the volatility of the FFR *can* increase. This observation hardly provides support for the hypothesis that the high asset valuation regimes were the product of low economic uncertainty, however.

Table 7 reports both means and standard deviations for the real interest rate and GDP growth, conditional on staying in each of the two regimes. The conditional steady state values are the means and standard deviations given by the VAR coefficients conditional on a particular regime being in place (see the Appendix for a precise definition.) For each statistic we report the median and 68% credible sets based on the posterior distribution of the MS-VAR parameters conditional on that regime. These statistics corroborate the non-steady state evidence presented above where the possibility of a regime shift is incorporated into expectations: the two regimes present a clear difference for the mean and volatility of the real interest rate. The high asset valuation regime is characterized by sharply lower expected real policy rates and lower uncertainty about those rates, while the opposite is true for the low asset valuation regime. By contrast, the high asset valuation regime is characterized by lower expected economic growth and higher economic uncertainty. Note that the conditional steady states do not depend on the estimated transition matrix so they show that the main conclusions on the differences across regimes are robust to estimation error on the transition matrix.

Figure 8 gives a visual impression of the result. The figure plots the "wealth ratio" (the inverse of cay^{MS} without removing the Markov-switching constant), along with the ten-year-ahead expected real federal funds rate, on separate scales. The red dashed line in the figure shows the most likely value of the unconditional mean of the wealth ratio in each regime (given by the inverse of the regime-probability weighted average of α_1 and α_2). The mean shows clear regime shifts in wealth ratios that move from high to low to high over the sample, coinciding with a low then high then low expected long-run real federal funds rate. Regime shifts in the expected federal funds rate are large, ranging from about 1% in the low expected interest rate regimes to 3% in the high expected interest rate regime.

These findings capture three distinct periods of post-WWII US economic history. The first manifestation of regime 2 is in the subperiod from 1952:Q1-1976:Q1 and coincides with the run-up of inflation in the 1960s and 1970s, accommodative monetary policy, and low real interest rates. Economists have provided several possible explanations for why monetary policy failed to react aggressively to inflation during those years. However, they generally tend to agree that this was a period of high uncertainty and possibly passive monetary policy (Clarida, Gali, and Gertler (2000); Lubik and Schorfheide (2004); Sims and Zha (2006); Bianchi (2013)). The occurrence of the low asset valuation regime, in the middle subperiod from 1976:Q2-2001:Q2, precedes by three years Volcker's appointment as Chairman of the Federal Reserve and the disinflation that followed. The first attempts to bring inflation down started in the late 1970s, but Volcker succeeded only in the early 1980s, perhaps because of the political backing provided by the Reagan administration (Bianchi and Ilut (2015)). As a result, the beginning of Great Moderation is generally placed in the mid-1980s, when the economy experienced a substantial reduction in volatility (McConnell and Perez-Quiros (2000); Stock and Watson (2002).) Macroeconomists interested in the Great Inflation tend to identify the change in the anti-inflationary stance of the Federal Reserve with the appointment of Volcker in August 1979. However, Sims and Zha (2006) estimate a structural MS-VAR and find a change in the conduct of monetary policy from less to more active toward the end of 1977, in line with the results here. Real interest rates increased significantly during the Volcker disinflation and they remained higher than in the 1970s for a prolonged period of time. In part this may have been attributable to a perceived need on the part of the Federal Reserve had to rebuild credibility for low and stable inflation.

The second occurrence of the high asset valuation regime in the subperiod 2001:Q3-2013:Q3 starts with the end of the information technology (IT) boom and the beginning of the Federal

Reserve's accommodative response to the recession that followed. Economists have identified the end of the Great Moderation with the 2008 recession, consistent with the estimated break patterns in Figure 7 for GDP growth uncertainty. At the same time, some have argued that monetary policy underwent a regime shift after the end of the IT boom (Campbell, Pflueger, and Viceira (2014)) and/or that interest rates were held "too low for too long"(Taylor (2007)) in response to the IT bust and the aftermath of 9/11. Asset values quickly recovered in 2002, and after a brief but dramatic decline in the financial crisis of 2007-2009, equity valuations resumed their upward march in 2009. This period of high equity valuations persists today with rates in the zero lower bound (ZLB) range coinciding with positive rates of inflation. Our estimates characterize this third subperiod as a return to a period of prolonged low real interest rates.

The three distinct *cay* regimes we estimate are remarkably close to the three distinct monetary policy regimes estimated by Campbell, Pflueger, and Viceira (2014), who use a completely different approach. Instead of identifying the break dates by using a cointegration relation in *cay*, they estimate break dates in the parameters of an estimated Taylor rule. Their first subperiod covers the period 1960:Q2-1977:Q1, the middle period is 1977:Q2-2000:Q4, and the last subperiod 2001:Q1 to the end of their sample 2011:Q4. They find that these regimes line up closely with shifts in estimated bond market betas. Although our focus is on regime shifts in an asset valuation ratio, *cay*, taken together, the results are suggestive of an important role for the Federal Reserve in driving persistent movements in equity and interest rate behavior.

5 Reaching for Yield?

We have found that low frequency swings in post-war asset valuation are strongly associated with low frequency shifts in the long-run expected value of interest rates, with low expected values for the real federal funds rate associated with high asset valuations, and vice versa. Moreover, while persistently low policy rates are associated with high asset valuations, this is not because they signal strong economic growth, favorable changes in inflation, or low uncertainty. This suggests that persistent changes in monetary policy affect asset valuations because they change the rate at which investor's discount future payouts, in a manner that is independent of uncertainty about the aggregate economy. This could occur simply because the Central Bank influences the riskless real interest rate, a component of the discount rate. But the magnitude of this discount rate effect would be amplified if it went beyond the riskless rate to affect risk *premia*. If a switch from a high to low interest rate regime prompts investors to take on more risk, to "reach for yield," then the risk premium component of the discount rate would fall, further stimulating risky asset values relative to economic fundamentals. The reverse would occur in a shift from persistently low expected rates to high. We refer to this general idea as the *reaching for yield* hypothesis, or RFY for brevity.[11]

Observe that a change in discount rates driven by the risk-free rate alone influences all assets in the same way, regardless of their riskiness. By contrast, RFY implies that investors shift portfolio allocations toward riskier/higher return assets in low interest rate environments. Thus a change in discount rates accompanied by RFY will have effects that differ across assets, depending on the riskiness of the asset. As interest rates move from high to low, RFY implies a greater increase in the market value, relative to fundamentals, of higher return/higher Sharpe ratio assets than it does for lower return/lower Sharpe ratio assets. Equivalently, risk premia should fall more for riskier assets. We investigate this possibility here, using data on book equity relative to the market values of individual stock market portfolios that exhibit strong cross-sectional variation in return premia.

The book-market ratios of individual assets can also be differentially affected by a shift in interest rates because their expected future earnings are differentially affected. This has nothing to do with risk premia or discount rates, so it is important to correct for possible differences in expected earnings growth when assessing the RFY.

To do so, we carry out a log-linearization that follows Vuolteenaho (2000) and constructs earnings from book-market and return data using clean surplus accounting. Let B_t denote book value and M_t denote market value, and let the logarithm of the book-market ratio $\log(B_t/M_t)$ be denoted θ_t. Vuolteenaho (2000) shows that the θ_t of an asset or portfolio can be decomposed as:

$$\theta_t = \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t r_{t+1+j} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t f_{t+1+j} - \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t e^*_{t+1+j} \tag{7}$$

where $\rho < 1$ is a parameter, and r_{t+1+j}, f_{t+1+j}, and e^*_{t+1+j} stand for log excess return, log risk-free rate, and log earnings, respectively.[12] In other words, the logarithm of the book-market ratio θ_t depends on the present discounted value (PDV) of expected excess returns (risk premia), expected risk-free rates, and expected earnings.

Given our objective to assess whether assets with different risk/return profiles respond differently to the regime changes identified above, we begin by focusing on the difference between the book-market ratios of portfolios known to have different risk/return profiles. Specifically, given two portfolios x and y, the spread in their book-market ratios, $\theta_{x,t} - \theta_{y,t}$, is given by:

$$\underbrace{\theta_{x,t} - \theta_{y,t}}_{\text{Spread in BM ratios}} = \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t \left(r_{x,t+1+j} - r_{y,t+1+j} \right)}_{\text{PDV of spread in risk premia}} - \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t \left(e^*_{x,t+1+j} - e^*_{y,t+1+j} \right)}_{\text{PDV of spread in expected earnings}}$$

Note that the risk-free rate has no affect on this spread, since all portfolios are affected in

[11] In what follows we use the terms "risk" premia and return premia interchangeably to refer to the expected return on an asset in excess of the risk-free rate. We remain agnostic as to whether the observed premia are attributable to mispricing, to covariance with systematic risk factors, or both.

[12] Specifically, e^* is the log of 1 plus the earnings-book ratio, adjusted for approximation error. See Vuolteenaho (2000).

the same way by the risk-free rate. Instead only the risk premium differential and expected earnings differential affect the book-market spread. Since RFR pertains only to the return premium differential, we adjust the book-market spread for the spread in expected earnings to isolate the return premium differential:

$$\underbrace{\theta_{x,t} - \theta_{y,t} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t \left(e^*_{x,t+1+j} - e^*_{y,t+1+j} \right)}_{\text{Spread in BM ratios adjusted for earnings}} = \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t \left(r_{x,t+1+j} - r_{y,t+1+j} \right)}_{\text{PDV of the spread in risk premia}} \tag{8}$$

The above expression shows that the spread in book-market ratios adjusted for expected future earnings is equal to the PDV of the spread in expected excess returns, or risk premia.

Denote the *adjusted* (for expected earnings) book-market ratio for portfolio x in regime r with a tilde as

$$\widetilde{\theta}^r_{x,t} \equiv \theta^r_{x,t} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t e^{*r}_{x,t+1+j}.$$

Let x denote a high return premia portfolio while y denotes a low return premia portfolio. Reaching for yield implies that, in a shift from a high ($r = 1$) to low ($r = 2$) interest rate regime, the adjusted book-market ratio of x should fall more than that of y, implying $\left(\widetilde{\theta}^1_{x,t} - \widetilde{\theta}^2_{x,t}\right) - \left(\widetilde{\theta}^1_{y,t} - \widetilde{\theta}^2_{y,t}\right) > 0$, or that the *difference-in-difference* of adjusted book-market ratios should be positive across regimes:

$$\left(\widetilde{\theta}^1_{x,t} - \widetilde{\theta}^1_{y,t}\right) - \left(\widetilde{\theta}^2_{x,t} - \widetilde{\theta}^2_{y,t}\right) > 0. \tag{9}$$

In summary, RFY implies that the *spread* in the *adjusted* book-market ratios between the high return/high risk portfolio and the low return/low risk portfolio should be greater in regime 1 than in regime 2.

To assess empirically whether the spread in adjusted book-market ratios between assets with different risk/return profiles is statistically different across the two regimes, we again estimate MS-VAR, now using portfolio data rather than macro data. The portfolio data are chosen to exhibit strong cross-sectional variation in average return premia and Sharpe ratios. The VAR specification takes the same form as equation (6), only the variables differ. Just as in the previous section, we impose the formerly estimated regime sequence for *cay* on the portfolio MS-VAR, but the parameters characterizing the different regimes, as well as the transition matrix, are freely estimated. The reasoning for doing so is the same as given above for the macro MS-VAR: we are interested in knowing whether the previously documented regime sequence for *cay* is characterized by evidence of RFY. This requires that we impose the previously estimated regime sequence, but since the MS-VAR parameters are freely estimated, the empirical procedure is free to find no evidence of structural change across these subsamples if indeed there is none.

We consider five different spread-portfolio VAR specifications using data on portfolios of stocks sorted by size (market capitalization) and book-market (BM) ratio, and portfolios of stocks sorted according to momentum (recent past return performance). The VARs differ according to which portfolios of a given size quintile are used to build the value book-market and return spreads. Specifically, the first portfolio VAR specification includes data on the five Fama-French risk factors $R_m - R_f$, SMB, HML, RMW, and CMA (Fama and French (2015)), the real FFR computed as FFR-inflation, and the following four variables:

1. Value BM spread: The difference between the logarithm of the BM ratio of the small (size quintile 1) high book-market portfolio and the logarithm of the BM ratio of the small (size 1) low book-market portfolio.

2. Momentum BM spread: The difference between the logarithm of the BM ratio of the extreme winner (M10) portfolio and the logarithm of the BM ratio of the extreme loser (M1) portfolio.

3. Value return spread: The difference between the excess return of the small (size 1) high BM portfolio and the excess return of the small (size 1) low BM portfolio.

4. Momentum return spread: The difference between the excess return of the extreme winner (M10) portfolio and the excess return of the extreme loser (M1) portfolio.

The other four portfolio VARs are obtained by replacing the value BM spread and the value return spread by the corresponding series for the portfolio in a different size quintile. We denote these size quintiles S1, S2,...,S5, where S1 is the smallest quintile and S5 the largest. The five Fama/French factors are always included in the VAR because they contain predictive information for the return premia on these portfolios, and thus for the PDV of the spread in expected excess returns. The BM data are constructed starting from the 25 Fama/French portfolios sorted by size and BM, and the 10 portfolios sorted on momentum.[13] The data on BM ratios for individual portfolios are constructed from CRSP and Compustat exactly as the Fama-French portfolio returns are constructed. We mimic the selection and breakpoints of this construction and compute the book-market ratio of each portfolio. The sample for this estimation is 1964:Q1-2013:Q4. This is shorter than the one previously used for cay^{MS} and the macro VAR because reliable data for book-market ratios are not available prior to 1964:Q1.

It is well known that equity assets formed by sorting stocks into portfolios on the basis of book-market and size, and on the basis of recent past return or earnings performance, exhibit sizable cross-sectional variation in return premia and conventional measures of the risk/return trade-off. High BM portfolios earn much higher average returns than low BM portfolios, especially in the small size categories. Along the momentum dimension, recent past winner stocks earn much higher returns than recent past losers. Table 8 reports sample statistics in our data

[13] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

on the annualized Sharpe ratios and means for the long-short value and momentum strategies. Specifically, the table reports these statistics for a portfolio that is long in the extreme value portfolio (highest BM ratio) and short in the extreme growth portfolio (lowest BM ratio) of a given size category, and for a portfolio that is long in the extreme winner portfolio (M10) and short in the extreme loser portfolio (M1). The table also reports the same statistics for the individual portfolio returns in excess of the risk-free rate, where the risk-free rate for this purpose is the one used by Fama-French, denoted R_f, and equal to the one-month Treasury bill rate. It is evident that the value strategies that go long in the high BM portfolio and short the low BM portfolio have high Sharpe ratios and return premia, especially those in the three smallest size categories. The Sharpe ratio for the smallest value long-short strategy is 0.62 with a mean return of 10%, while that on the medium size quintile has a Sharpe ratio of 0.42 and mean return of 7%. The momentum strategy has an annualized Sharpe ratio of 0.64 and mean return of over 15%. Moreover, for the individual value, growth, winner, loser portfolios, value portfolios have much higher risk-premia than growth portfolios in the same size category, while the winner portfolio has a much higher risk premium than the loser portfolio. Indeed, the loser portfolio has a negative average return premium in the full sample, suggesting that it is not a risky asset and may even provide insurance. These statistics confirm the well known heterogeneity in the risk/return profiles of these portfolios.

Our objective is exploit this heterogeneity to isolate the effects of the previously estimated regime changes in the mean of *cay* on the adjusted BM ratios of different portfolios. To do so, we begin by computing the conditional steady states of the adjusted BM *spreads* between the high and low return premia portfolios, $\widetilde{\theta}^r_{t,x} - \widetilde{\theta}^r_{t,y}$, for each regime. As above, the conditional steady state values are the means and standard deviations given by the VAR coefficients conditional on a particular regime being in place and we report the median and 68% credible sets based on the posterior distribution of the MS-VAR parameters conditional on that regime. The high and low return premia portfolios along the BM dimension are always the extreme value (highest BM) and the extreme growth portfolio (lowest BM), respectively, in each size category. Likewise, the high and low return premia portfolios along the momentum dimension are always the extreme winner (M10) and extreme loser portfolio (M1). These results are reported in the first two rows of Table 9. The third row reports the the difference-in-difference of adjusted book-market ratios between the high and low return premia portfolios across the two regimes, i.e., the difference between the spreads $\left(\widetilde{\theta}^1_{x,t} - \widetilde{\theta}^1_{y,t}\right) - \left(\widetilde{\theta}^2_{x,t} - \widetilde{\theta}^2_{y,t}\right)$, as implied by the VAR estimates. To interpret the table, keep in mind that regime 1 is the low asset valuation/high interest rate regime, while regime 2 is the high asset valuation/low interest rate regime.

For all five size quintiles and every VAR, Table 9 shows that the adjusted BM spreads between the high and low return premia portfolios are positive in both regimes. This is not surprising because portfolios that have higher risk premia should have lower market values, holding fixed expected earnings and book value. Importantly, however, the table shows that these spreads are greater in regime 1 (low asset valuations/high interest rates) than in regime 2 (high asset valuations/low interest rates). Thus the difference-in-difference across regimes is always positive. This implies that the adjusted book-market ratios of high return/high risk portfolios fall more in a shift from high to low interest rate regime than do those of low return/low risk portfolios. Put differently, the return premia of evidently riskier/higher return assets decline more in environments with persistently high aggregate wealth ratios and low Federal Reserve policy rates than do less risky/lower return assets.

The third row also reports the 68% posterior credible intervals in parentheses for the difference-in-difference. The break in the BM spreads is proportionally smaller than the break in the momentum spreads. However, in all cases it is positive and, with the exception of the adjusted BM spread between the value and growth stocks in the fourth largest size quintile (size 4), this difference is strongly statistically significant. These results are supportive of a channel that implies an increased appetite for risk-taking in low interest rate environments.

The magnitude of the RFY channel for some portfolios is striking, and the case of momentum in particular deserves emphasis. The results indicate that, in every VAR estimated, the spread in adjusted BM ratios between the winner and loser portfolios is nearly one and a half times as high in the high interest rate regime than in the low interest rate regime, suggesting quantitatively large shifts toward greater risk-taking in the low interest rate subperiods. It is notable that momentum investing, with portfolios that exhibit the largest shift in these spreads that we find, is known to be among the most volatile equity investment strategies studied, one that is subject to infrequent but extreme crashes especially prevalent in times of crisis or panic (Daniel and Moskowitz (2013)).

These findings may be equivalently stated in terms of risk premia (see 8): long-short portfolios that exploit value and momentum spreads have lower risk premia in low interest rate regimes, and higher risk premia in high interest rate regimes. Specifically, the PDV of all future risk premia spreads is estimated to be considerably lower in low interest regimes than in high interest rate regimes.

The findings in Table 9 report the PDV of risk premia, conditional on one regime or another. We can also estimate the PDV of risk premia as it evolves over the sample, rather than in conditional steady state. These values are estimated as the VAR forecasts of the return premia spreads, conditional on information in the VAR data as of time t and taking into account the probability of a regime change as of t. (The Appendix explains how this is computed from the MS-VAR estimates.) The first row of Figure 9 reports these conditional forecasts for the value-growth spreads in each size category and the winner-loser spread. (The results for the latter are arbitrarily reported using the estimates of the (S5) VAR, since the results are very similar across the five VAR specifications).[14] Although the risk premia are volatile, there are some clear patterns. With the exception of the S4 value-growth portfolio, risk premia fluctuate around a lower point in the high asset valuation/low interest rate regime than they do in the low asset valuation/high interest rate regime. For the four long-short strategies other than this one, the estimated risk premia reach lows or near-lows in the post-millennial period, during the second occurrence of the high asset valuation/low interest rate regime, after shooting up briefly in the aftermath of the financial crisis of 2007-2008. The risk premia then return to low levels in the post-crisis ZLB period.

We also investigate how the risk premia on the individual value and momentum portfolios, as opposed to the long-short strategies, have changed over the sample. To do so, we form an estimate of the first term on the right-hand-side of (7), namely the PDV of all future risk premia, for each portfolio. These estimates are formed from six MS-VAR specifications. The first five VARs pertain to portfolios sorted according to size and book-market, and differ by size quintile. The sixth VAR uses the same data but for the momentum portfolios. Specifically, each VAR includes data on the five Fama-French risk factors $R_m - R_f$, SMB, HML, RMW, and CMA (Fama and French (2015)), the real FFR computed as FFR-inflation, and the following four variables:

1. The logarithm of the BM ratio of the high book-market portfolio (value) in a given size quintile (small, S2, S3, S4, large) or the logarithm of the BM ratio of the extreme winner (M10) portfolio.

2. The logarithm of the BM ratio of the low book-market portfolio (growth) in a given size quintile (small, S2, S3, S4, Large) or the logarithm of the BM ratio of the extreme loser (M1) portfolio.

3. The excess return of the BM ratio of the high book-market portfolio (value) in a given size quintile (small, S2, S3, S4, large) or the excess return of the BM ratio of the extreme winner (M10) portfolio.

4. The excess return of the BM ratio of the low book-market portfolio (growth) in a given size quintile (small, S2, S3, S4, Large) or the excess return of the BM ratio of the extreme loser (M1) portfolio.

The specification of these six VARs differs from that of the five VARs used above to estimate adjusted BM spreads across portfolios, where the momentum variables were included in each size/book-market VAR. There are two reasons we use a different VAR specification for

[14]Both the estimated risk premia over time and the conditional steady states are computed for each draw of the VAR parameters by taking into account the possibility of regime changes. The lines in the figure correspond to those for the median values across draws.

estimating the risk premia on the spreads than for estimating those for the individual portfolios. First, although the VAR specified above for the individual portfolios could in principal be used to estimate risk premia for the long-short spreads, in practice risk premia for the spreads are estimated far more precisely using a VAR that includes data on the spreads, rather than data on the individual portfolios separately. Second, including the momentum variables in the book-market VARs for the individual portfolios (as was done for the spreads) would have required two additional variables in each VAR compared to the case for the spreads, a number that overwhelmed the numerical calculation and caused many draws to be rejected due to non-stationarity. Stationarity is required to form accurate estimates of the conditional expectations embedded in risk premia and the PDV of expected returns. The above specification with six VARs exhibited no problems with stationarity.

The results on risk premia for the individual portfolios are displayed in the second and third rows of Figure 9. The solid (blue) line is the estimated PDV of risk premia over time, while the dashed (red) line is the value conditional on the regime in place. For nine of the twelve portfolios, the PDV of risk premia are all lower in the high asset valuation/low interest rate regime than in the low asset valuation/high interest rate regime. For several portfolios these premia reach lows or near-lows in the post-2000 period and the post-crisis period, when in the latter case interest rates entered the ZLB range. Almost all portfolios exhibit an increase in risk premia during the years corresponding to the financial crisis, but in every case risk premia decline subsequent to the crisis. It is worth noting that the one big exception to this is the loser portfolio, which as noted has an average return over the risk-free rate that is *negative*, suggesting not only that it is not a risky asset, but may even provide insurance. If so, we would not expect the behavior of this portfolio to be even qualitatively consistent with those of the risky portfolios, since the predictions of either literature discussed in the introduction for how the risk premium on an asset changes with the interest rate are predicated on the presumption that the asset in question carries a positive risk premium on average.

The findings reported in Figure 9 are point estimates based on the median draw from the estimated posterior distribution for the PDV of risk premia. But it is also possible to make inferences on the extent to which risk premia are lower in high asset valuation/low interest rate regimes by examining other moments of the posterior distribution. For example, we can use the estimated posterior distribution to compute the exact probability that a high asset valuation/low interest rate regime is characterized by lower risk premia than the low asset valuation/high interest rate regime. These probabilities are computed as the percentage of draws from the posterior distribution for which a decline in risk premia occurs when moving from regime 1 to regime 2. The results are reported in Table 10. The first and second rows refer to the risk premia for the long-short spread portfolios, while the third and fourth rows refer to the risk premia on individual portfolios.

Table 10 shows that, for five out of the six spread portfolios, the probability that the difference in risk premia is positive (i.e., that premia go down when entering a high asset valuation/low interest rate regime) is at least 85%. The exception is value-growth size 4 spread portfolio. But three of the value-growth portfolios assign probabilities in excess of 90%, and the winner-loser portfolios assign a probability of effectively 100%. For the individual portfolios, the estimation likewise assigns a large probability of a decline in risk premia in most cases. For 8 out of 12 portfolios this probability is at least 85%. Moreover, taking the results for all portfolios together, the estimated probability appears related to the riskiness of the portfolio in the expected way. The long-short (spread) portfolios are arguably the riskiest since they require leverage. For these portfolios the probability that risk premia fall in the high asset valuation/low interest rate regime is above 90% for the three value-growth strategies with the highest average risk premia (those for S1, S2, and S3), and is often close to 100% for winner-loser strategies. Similarly, the high risk premia value and winner individual portfolios exhibit greater probabilities of a decline in premia than do the lower risk premia growth and loser portfolios. Finally, in line with the evidence presented above, the evidence says that the probability of a decline in risk premia for the loser portfolio is close to zero (5%), consistent with the evidence that this portfolio does not command a positive risk premium on average and so would not be expected to exhibit a decline in its premium in regimes where the appetite for risk-taking is on the rise.

In summary, the results from estimating portfolio VARs indicate that, low interest rate regimes are associated with lower risk premia for most portfolios that carry a positive (average) risk premium, while the riskier portfolios exhibit larger declines in premia over the course of a low interest rate regime than do less risky portfolios. The findings are supportive of reaching for yield theories. They present a challenge, however, for theories that explain persistently low interest rate environments with shocks that shift in the composition of wealth toward more risk averse investors (e.g., Barro and Mollerus (2014); Caballero and Farhi (2014); Hall (2016)). In contrast to RFY, these theories imply that low interest rates should coincide with higher risk premia rather than lower. The findings here indicate that, not only are risk premia lower conditional on being in a low interest rate regime but, for most portfolios the estimated historical variation in these premia indicates that they declined early in the pre-2000 period and again at the onset of the ZLB period, after a brief but sharp spike upward during the financial crisis.

6 Conclusion

This paper presents evidence of infrequent shifts, or "breaks," in the mean of the consumption-wealth variable cay_t, an asset market valuation ratio driven by fluctuations in stock market wealth relative to economic fundamentals. These infrequent mean shifts generate low frequency fluctuations in asset values relative to fundamentals as measured by cay. A Markov-switching cay_t, denoted cay_t^{MS}, is estimated and shown to be less persistent and have superior forecasting power for excess stock market returns compared to the conventional estimate. Evidence from a Markov-Switching VAR shows that these low frequency swings in post-war asset valuation are strongly associated with low frequency swings in the long-run expected value of the Federal Reserve's primary policy interest rate, with low expected values for the real federal funds rate associated with high asset valuations, and vice versa. The findings suggest that the expectation of persistently low policy rates may be partly responsible for the high asset valuations of the last several years, and vice versa for the low asset valuation regime in the middle part of our post-war sample.

At the same time, we find no evidence that the estimated structural shifts to high asset valuation regimes and persistently low policy rates are associated with optimism about the future in the form of expectations for stronger long-run economic growth or lower uncertainty about that growth. Instead, the results suggest that infrequent regime shifts to high asset valuations are driven by persistent shifts in the stance of monetary policy that merely reduce the rate at which investors discount future payouts, without engendering favorable expectations for real economic growth, inflation, or uncertainty. Moreover, this finding does not appear to be attributable to a movement in discount rates driven solely by the expected real interest rate. Instead, we find evidence of low frequency shifts in the risk premium component of the discount rate, consistent with the hypothesis that investors' willingness to tolerate risk in equity markets rises whenever the long-run expected value of the real federal funds rate is low. We find that the magnitude of this effect is especially pronounced for some of the most volatile equity investment strategies subject to infrequent but extreme crashes, such as those based on leveraged momentum investing.

The theoretical literature on such "reaching for yield" channels is still in its infancy, while this paper is an empirical study. More theoretical work is needed to understand how and why this channel may arise, and to elicit additional implications for asset markets and macroeconomic quantities.

Appendix

Data Appendix

This appendix describes the data used in this study.

CONSUMPTION

Consumption is measured as either total personal consumption expenditure or expenditure on nondurables and services, excluding shoes and clothing. The quarterly data are seasonally adjusted at annual rates, in billions of chain-weighted 2005 dollars. The components are chain-weighted together, and this series is scaled up so that the sample mean matches the sample mean of total personal consumption expenditures. Our source is the U.S. Department of Commerce, Bureau of Economic Analysis.

LABOR INCOME

Labor income is defined as wages and salaries + transfer payments + employer contributions for employee pensions and insurance - employee contributions for social insurance - taxes. Taxes are defined as [wages and salaries/(wages and salaries + proprietors' income with IVA and CCADJ + rental income + personal dividends + personal interest income)] times personal current taxes, where IVA is inventory valuation and CCADJ is capital consumption adjustments. The quarterly data are in current dollars. Our source is the Bureau of Economic Analysis.

POPULATION

A measure of population is created by dividing real total disposable income by real per capita disposable income. Our source is the Bureau of Economic Analysis.

WEALTH

Total wealth is household net worth in billions of current dollars, measured at the end of the period. A break-down of net worth into its major components is given in the table below. Stock market wealth includes direct household holdings, mutual fund holdings, holdings of private and public pension plans, personal trusts, and insurance companies. Nonstock wealth includes tangible/real estate wealth, nonstock financial assets (all deposits, open market paper, U.S. Treasuries and Agency securities, municipal securities, corporate and foreign bonds and mortgages), and also includes ownership of privately traded companies in noncorporate equity, and other. Subtracted off are liabilities, including mortgage loans and loans made under home equity lines of credit and secured by junior liens, installment consumer debt and other. Wealth is measured at the end of the period. A timing convention for wealth is needed because the level of consumption is a flow during the quarter rather than a point-in-time estimate as is wealth (consumption data are time-averaged). If we think of a given quarter's consumption data as measuring spending at the beginning of the quarter, then wealth for the quarter should be measured at the beginning of the period. If we think of the consumption data as measuring spending at the end of the quarter, then wealth for the quarter should be measured at the end of the period. None of our main findings discussed below (estimates of the cointegrating parameters, error-correction specification, or permanent-transitory decomposition) are sensitive to this timing convention. Given our finding that most of the variation in wealth is not associated with consumption, this timing convention is conservative in that the use of end-of-period wealth produces a higher contemporaneous correlation between consumption growth and wealth growth. Our source is the Board of Governors of the Federal Reserve System. A complete description of these data may be found at http://www.federalreserve.gov/releases/Z1/Current/.

STOCK PRICE, RETURN, DIVIDENDS

The stock price is measured using the Center for Research on Securities Pricing (CRSP) value-weighted stock market index covering stocks on the NASDAQ, AMEX, and NYSE. The data are monthly. The stock market price is the price of a portfolio that does not reinvest dividends. The CRSP dataset consists of $vwretx(t) = (P_t/P_{t-1}) - 1$, the return on a portfolio that doesn't pay dividends, and $vwretd_t = (P_t + D_t)/P_t - 1$, the return on a portfolio that does pay dividends. The stock price index we use is the price P_t^x of a portfolio that does not reinvest dividends, which can be computed iteratively as

$$P_{t+1}^x = P_t^x \left(1 + vwretx_{t+1}\right),$$

where $P_0^x = 1$. Dividends on this portfolio that does not reinvest are computed as

$$D_t = P_{t-1}^x \left(vwretd_t - vwretx_t\right).$$

The above give monthly returns, dividends and prices. The annual log return is the sum of the 12 monthly log returns over the year. We create annual log dividend growth rates by summing the log differences over the 12 months in the year: $d_{t+12} - d_t = d_{t+12} - d_{t+11} + d_{t+11} - d_{t+10} + \cdots + d_{t+1} - d_t$. The annual log price-dividend ratio is created by summing dividends in levels over the year to obtain an annual dividend in levels, D_t^A, where t denotes a year hear. The annual observation on P_t^x is taken to be the last monthly price observation of the year, P_t^{Ax}. The annual log price-dividend ratio is $\ln\left(P_t^{Ax}/D_t^A\right)$.

PRICE DEFLATOR

The nominal after-tax labor income and wealth data are deflated by the personal consumption expenditure chain-type deflator (2005=100), seasonally adjusted. In principle, one would like a measure of the price deflator for total flow consumption here. Since this variable is unobservable, we use the total expenditure deflator as a proxy. Our source is the Bureau of Economic Analysis.

Gibbs Sampling Algorithm

This appendix describes the Bayesian methods used to characterize uncertainty in the parameters of the regression (2). To simplify notation, we denote the vector containing all variables whose coefficients are allowed to vary over time $x_{M,t}$, while $x_{F,t}$ is used to denote the vector containing all the variables whose coefficients are kept constant. We then obtain:

$$c_t = \alpha_{\xi_t^\alpha} x_{M,t} + \beta x_{F,t} + \sigma \varepsilon_t$$

where, in our case, $\beta = \left[\beta_a, \beta_y, b_{a,-k}, ..., b_{a,+k}, b_{y,-k}, ..., b_{y,+k}\right]$ and the vector $x_{M,t}$ is unidimensional and always equal to 1.

Suppose the Gibbs sampling algorithm has reached the $r-th$ iteration. We then have draws for β_r, $\alpha_{\xi_t^\alpha,r}$, σ_r, H_r^α, and $\xi_r^{\alpha,T}$, where $\xi_r^{\alpha,T} = \{\xi_{1,r}^\alpha, \xi_{2,r}^\alpha, ..., \xi_{T,r}^\alpha\}$ denotes a draw for the whole regime sequence. The sampling algorithm is described as follows.

1. **Sampling** β_{r+1}: Given $\alpha_{\xi_t^\alpha,r}$, σ_r, and $\xi_r^{\alpha,T}$ we transform the data:

$$\widetilde{c}_t = \frac{c_t - \alpha_{\xi_t^\alpha,r} x_{M,t}}{\sigma_r} = \beta \frac{x_{F,t}}{\sigma_r} + \varepsilon_t = \beta \widetilde{x}_t + \varepsilon_t.$$

The above is a regression with fixed coefficients β and standardized residual shocks. Standard Bayesian methods may be used to draw the coefficients of the regression. We assume a Normal conjugate prior $\beta \sim N\left(B_{\beta,0}, V_{\beta,0}\right)$), so that the conditional (on $\alpha_{\xi_t^\alpha,r}$, σ_r, and $\xi_r^{\alpha,T}$) posterior distribution is given by

$$\beta_{r+1} \sim N\left(B_{\beta,T}, V_{\beta,T}\right)$$

with $V_{\beta,T} = \left(V_{\beta,0}^{-1} + \widetilde{X}_F' \widetilde{X}_F\right)^{-1}$ and $B_{\beta,T} = V_{\beta,T}\left[V_{\beta,0}^{-1} B_{\beta,0} + \widetilde{X}_F' \widetilde{C}\right]$, where $\widetilde{C}$ and $\widetilde{X}_F$ collect all the observations for the transformed data and $B_{\beta,0}$ and $V_{\beta,0}^{-1}$ control the priors for the fixed coefficients of the regression. With flat priors, $B_{\beta,0} = 0$ and $V_{\beta,0}^{-1} = 0$ and $B_{\beta,T}$ and $V_{\beta,T}$ coincide with the maximum likelihood estimates, conditional on the other parameters.

2. **Sampling** $\alpha_{i,r+1}$ for $i = 1, 2$: Given β_{r+1}, σ_r, and $\xi_r^{\alpha,T}$ we transform the data:

$$\widetilde{c}_t = \frac{c_t - \beta_{r+1} x_{F,t}}{\sigma_r} = \alpha_{\xi_t^\alpha} \frac{x_{M,t}}{\sigma_r} + \varepsilon_t = \alpha_{\xi_t^\alpha} \widetilde{x}_{M,t} + \varepsilon_t.$$

The above regression has standardized shocks and Markov-switching coefficients in the transformed data. Using $\xi_r^{\alpha,T}$ we can group all the observations that pertain to the same regime i. Given the prior $\alpha_i \sim N\left(B_{\alpha_i,0}, V_{\alpha_i,0}\right)$) for $i = 1, 2$ we use standard Bayesian methods to draw α_i from the conditional (on β_{r+1}, σ_r, and $\xi_r^{\alpha,T}$) posterior distribution:

$$\alpha_{i,r+1} \sim N\left(B_{\alpha_i,T}, V_{\alpha_i,T}\right) \text{ for } i = 1, 2$$

where $V_{\alpha_i,T} = \left(V_{\alpha_i,0}^{-1} + \widetilde{X}_{M,i}'\widetilde{X}_{M,i}\right)^{-1}$ and $B_{\alpha_i,T} = V_{\alpha_i,T}\left[V_{\alpha_i,0}^{-1}B_{\alpha_i,0} + \widetilde{X}_{M,i}'\widetilde{C}_i\right]$ where $\widetilde{C}_i$ and $\widetilde{X}_{M,i}$ collect all the observations for the transformed data for which regime i is in place. The parameters $B_{\alpha_i,0}$ and $V_{\alpha_i,0}^{-1}$ control the priors for the MS coefficients of the regression: $\alpha_i \sim N\left(B_{\alpha_i,0}, V_{\alpha_i,0}\right)$ for $i = 1, 2$. With flat priors, we have $B_{\alpha_i,0} = 0$ and $V_{\alpha_i,0}^{-1} = 0$ and $B_{\alpha_i,T}$ and $V_{\alpha_i,T}$ coincide with the maximum likelihood estimates, conditional on the other parameters.

3. **Sampling** σ_{r+1}: Given β_{r+1}, $\alpha_{\xi_t^\alpha,r+1}$, and $\xi_r^{\alpha,T}$ we can compute the residuals of the regression:

$$\widetilde{c}_t = c_t - \beta_{r+1}x_{F,t} - \alpha_{\xi_t^\alpha}x_{M,t} = \sigma\varepsilon_t.$$

With the prior that σ has an inverse gamma distribution, $\sigma \sim IG\left(Q_0, v_0\right)$, we use Bayesian methods to draw σ_{r+1} from the conditional (on β_{r+1}, $\alpha_{\xi_t^\alpha,r+1}$, and $\xi_r^{\alpha,T}$) posterior inverse gamma distribution:

$$\sigma_{r+1} \sim IG\left(Q_T, v_T\right), \; v_T = T + v_0, \; Q_T = Q_0 + E'E$$

where E is a vector containing the residuals, T is the sample size, and Q_0 and v_0 control the priors for the standard deviation of the innovations: $\sigma \sim IG\left(Q_0, v_0\right)$. With flat priors, we have $Q_0 = 0$ and $v_0 = 0$. The mean of a random variable with distribution $\sigma \sim IG\left(Q_T, v_T\right)$ is Q_T/v_T. With flat priors we have $Q_0 = 0$ and $v_0 = 0$, and the mean of σ is therefore $\left(E'E\right)/T$, which coincides with the standard maximum likelihood (MLE) estimate of σ, conditional on the other parameters.

4. **Sampling** $\xi_{r+1}^{\alpha,T}$: Given β_{r+1}, $\alpha_{\xi_t^\alpha,r+1}$, and H_r^α we can obtain filtered probabilities for the regimes, as described in Hamilton (1994). Following Kim and Nelson (1999) we then use a Multi-Move Gibbs sampling to draw a regime sequence $\xi_{r+1}^{\alpha,T}$.

5. **Sampling** H_{r+1}^α: Given the draws for the MS state variables $\xi_{r+1}^{\alpha,T}$, the posterior for the transition probabilities does not depend on other parameters of the model and follows a Dirichlet distribution if we assume a prior Dirichlet distribution.[15] For each column of H_{r+1}^α the posterior distribution is given by:

$$H_{r+1}^\alpha(:, i) \sim D(a_{ii}^\alpha + \eta_{ii,r+1}^\alpha, a_{ij}^\alpha + \eta_{ij,r+1}^\alpha)$$

where $\eta_{ij,r+1}^\alpha$ denotes the number of transitions from state i^α to state j^α based on $\xi_{r+1}^{\alpha,T}$, while a_{ii}^α and a_{ij}^α the corresponding priors. With flat priors, we have $a_{ii}^\alpha = 0$ and $a_{ij}^\alpha = 0$.

6. If $r + 1 < R$, where R is the desired number of draws, go to step 1, otherwise stop.

[15] The Dirichlet distribution is a generalization of the beta distribution that allows one to potentially consider more than 2 regimes. See e.g., Sims and Zha (2006).

These steps are repeated until convergence to the posterior distribution is reached. We check convergence by using the Raftery-Lewis Diagnostics for each parameter in the chain. See section below. We use the draws obtained with the Gibbs sampling algorithm to characterize parameter uncertainty in Table 1.

6.1 Computing cay^{MS}

Our estimate of cay_t^{MS} is based on the posterior mode of the parameter vector $\boldsymbol{\theta}$ and the corresponding regime probabilities. The filtered probabilities reflect the probability of a regime conditional on the data up to time t, $\pi_{t|t} = p(\xi_t^\alpha | Y^t; \boldsymbol{\theta})$, for $t = 1, ..., T$, and are part of the output obtained computing the likelihood function associated with the parameter draw $\boldsymbol{\theta} = \left\{\beta, \alpha_{\xi_t^\alpha}, \sigma, \boldsymbol{H}^\alpha\right\}$. They can be obtained using the following recursive algorithm:

$$\begin{aligned} \pi_{t|t} &= \frac{\pi_{t|t-1} \odot \eta_t}{\mathbf{1}' \left(\pi_{t|t-1} \odot \eta_t\right)} \\ \pi_{t+1|t} &= \boldsymbol{H}^\alpha \pi_{t|t} \end{aligned}$$

where η_t is a vector whose j-th element contains the conditional density $p(c_t | \xi_t^\alpha = j, x_{M,t}, x_{F,t}; \boldsymbol{\theta})$, i.e.,

$$p(c_t | \xi_t^\alpha = j, x_{M,t}, x_{F,t}; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\left\{c_t - (\alpha_j x_{M,t} + \beta x_{F,t})\right\}^2}{2\sigma^2} \right),$$

the symbol $\odot$ denotes element by element multiplication, and $\mathbf{1}$ is a vector with all elements equal to 1. To initialize the recursive calculation we need an assumption on the distribution of ξ_0^α. We assume that the two regimes have equal probabilities: $p(\xi_0^\alpha = 1) = .5 = p(\xi_0^\alpha = 2)$.

The smoothed probabilities reflect all the information that can be extracted from the whole data sample, $\pi_{t|T} = p(\xi_t^\alpha | Y^T; \boldsymbol{\theta})$. The final term, $\pi_{T|T}$ is returned with the final step of the filtering algorithm. Then, a recursive algorithm can be implemented to derive the other probabilities:

$$\pi_{t|T} = \pi_{t|t} \odot \left[\boldsymbol{H}^{\alpha\prime} \left(\pi_{t+1|T} \left(\div \right) \pi_{t+1|t} \right) \right]$$

where $(\div)$ denotes element by element division.

In using the DLS regression (2) to estimate cointegrating parameters, we lose 6 leads and 6 lags. For estimates of cay_t^{FC}, we take the estimated coefficients and we apply them to the whole sample. To extend our estimates of cay_t^{MS} over the full sample, we can likewise apply the parameter estimates to the whole sample but we need an estimate of the regime probabilities in the first 6 and last 6 observations of the full sample. For this we run the Hamilton filter from period from -5 to $T + 6$ as follows. When starting at -5, we assume no lagged values are available and the DLS regression omits all lags, but all leads are included. When at $t = -4$ we assume only one lag is available and the DLS regression includes only one lag, and so on, until we reach $t = 0$ when all lags are included. Proceeding forward when $t = T + 1$ is reached we assume all lags are available and all leads except one are available, when $t = T + 2$, we assume all lags and all leads but two are available, etc. Smoothed probabilities are then computed with standard methods as they only involve the filtered probabilities and the transition matrix $\boldsymbol{H}^{\alpha}$.

Estimation of Fractionally Integrated Models

In order to evaluate the likelihood for the fractionally integrated model we closely follow Muller and Watson (2013). We in fact use a series of Matlab codes that are available on Mark Watson's webpage. The first step consists of computing the cosine transformation of cay:

$$f_j = \iota_{jT} T^{-1} \sum_{t=1}^{T} \sqrt{2} \cos\left(j(t - 0.5)\pi T^{-1}\right) cay_t \quad \text{for } j = 1, ..., k.$$

where $\iota_{jT} = (2T/(j\pi)) \sin(j\pi/(2T))$. As explained in Muller and Watson (2013), this transformation is useful to isolate variation in the sample at different frequencies. Specifically, f_j captures variation at frequency $j\pi/T$. Mueller and Watson (2008, 2013) explain that working with a subset of the cosine transformations implies truncating the information set. They provide two reasons for why this is a convenient approach. First, given that each variable is a weighted average of the original data, a central limit allows to work with a limiting Gaussian distribution. Second, such a choice implies robustness of the results: Low-frequency information is used to study the low-frequency properties of the model. Given that we are mostly interested in the low frequency properties of cay, we can work using a limited number of (low) frequencies. We therefore choose $k = 12$.

We can then collect all the cosine transformations in a vector $X_{T,1:k}$ and compute an invariant transformation $X^s_{T,1:k} = X_{T,1:k}/\sqrt{X'_{T,1:k} X_{T,1:k}}$ (notice that this implies that the results that will follow are independent of scale factors). As explained in Muller and Watson (2013), the limiting density for the invariant transformations is given by:

$$p_{X^S}(x^s) = \frac{1}{2}\Gamma(k/2)\, \pi^{-k/2} \left|\Sigma_X\right|^{-1/2} \left(x^{s\prime} \Sigma_X^{-1} x^s\right)^{-q/2} \tag{10}$$

where $X^s = X_{1:k}/\sqrt{X'_{1:k} X_{1:k}}$, $\Sigma_X = E\left(X^s X^{s\prime}\right)$, and Γ is the gamma function.

We then assume a fractionally integrated model for cay_t: $(1 - L)^d cay_t = u_t$, where L is the lag operator and u_t is an $I(0)$ process and d is a parameter that is allowed to be fractional. The fractional model implies a binomial series expansion in the lag operator:

$$\begin{aligned}
(1 - L)^d cay_t &= \left[\sum_{k=0}^{\infty} \binom{d}{k} (-L)^k\right] cay_t \\
&= \left[\sum_{k=0}^{\infty} \frac{\prod_{a=0}^{k-1} (d - a)\,(-L)^k}{k!}\right] cay_t \\
&= \left[1 - dL + \frac{d\,(d-1)}{2!} L^2 - ...\right] cay_t
\end{aligned}$$

Note that when $d = 1$, the fractional integrated model implies that cay_t has a unit root, $cay_t = cay_{t-1} + u_t$, while for $d = 0$, $cay_t = u_t$, i.e. cay_t is an $I(0)$ process.

We compute the covariance matrix $\Sigma_X(d)$ associated with different values of d in the fractionally integrated model. The matrix $\Sigma_X(d)$ is obtained in two steps. First, we compute the matrix of autocovariances $\Sigma(d)$ associated with a fractionally integrated model. The $(i, i+h)$ element of this matrix is given by the autocovariance $\gamma(h)$:

$$\Sigma(d)_{(i,i+h)} = \gamma(h) = \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} \frac{\Gamma(h+d)}{\Gamma(1+h-d)}$$

Second, we transform the autocovariance matrix $\Sigma(d)$ in order to obtain the covariance matrix for the cosine transformations: $\Sigma_X(d) = \Psi' \Sigma(d) \Psi$ where Ψ is a $(T \times k)$ matrix collecting all the weights used for the cosine transformation:

$$\Psi_{(t,j)} = \iota_{jT} T^{-1} \sum_{t=1}^{T} \sqrt{2} \cos\left(j(t-0.5)\pi T^{-1}\right)$$

Finally, we evaluate (10) to obtain the likelihood for the different values of d given that $\Sigma_X(d)$ is now a function of the parameter d of the fractionally integrated model.

Robustness: Markov-Switching in Other Parameters

We here analyze two alternative models and compare them to our benchmark model in which only the constant is allowed to change over time and the fixed coefficient regression. In the first model, we allow for heteroskedasticity and changes in the constant. In the second model, we only allow for heteroskedasticity. We then use the Bayesian information criterion (BIC) to compare the different models. This is computed as:

$$BIC = -2(maxli) + k \log(T/(2\pi))$$

where $maxli$ is the maximized likelihood, k is the number of parameters, and T the sample size. Therefore, The Bayesian information criterion automatically penalizes models that have more parameters.

Table A.1 reports the estimates for the key parameters and the BIC for each model. We find that the BIC is minimized by the model is the one that allows for both heteroskedasticity and changes in the constant (MS α and MS σ). Our benchmark model with only changes in the constant (MS α only) is preferred to the model that only allows for heteroskedasticity (MS σ only) and the fixed coefficient regression (FC). Therefore, our results clearly support the hypothesis of shifts in the constant. Furthermore, the estimates for the cointegrating vector are basically unchanged when introducing heteroskedasticity in our benchmark model. For this reason, we choose the simpler model with only shifts in the constant as our benchmark model.

Model	α_1	α_2	β_a	β_y	σ_1	σ_2	BIC
MS α and MS σ	0.9186	0.8810	0.2599	0.6162	0.0016	0.0105	−1472.0
MS α only	0.9186	0.8808	0.2606	0.6156	0.0080		−1443.7
MS σ only	0.8056		0.1275	0.7845	0.0029	0.0204	−1281.0
FC	0.8706		0.1246	0.7815	0.0158		−1173.5

Table A.1. The table reports the estimates for the cointegration parameters, the estimates for the volatilities, and the Bayesian Information Criterion (BIC) for four different models. The BIC is used to compare the fit of different models taking into account the number of parameters used in the estimates. MS α and MS σ: The model allows for changes in the constant and heteroskedasticity. MS α only: Benchmark model with only changes in the constant. MS σ only: The model allows for heteroskedasticity, but not changes in the constant. FC: Standard fixed coefficient regression.

Additional Statistical Results

The tables below pertain to convergence of the Gibbs sampling algorithm.

Variable	Total(N)	I-stat	Variable	Total(N)	I-stat	Variable	Total(N)	I-stat
α_1	17413	9.541	Δa_{t+1}	1799	0.986	Δy_{t-4}	1850	1.014
α_2	16949	9.287	Δy_{t+1}	1812	0.993	Δa_{t+4}	1793	0.982
β_a	1918	1.051	Δa_{t-2}	1830	1.003	Δy_{t+4}	1820	0.997
β_y	1843	1.01	Δy_{t-2}	1801	0.987	Δa_{t-5}	1797	0.985
σ	1797	0.985	Δa_{t+2}	1886	1.033	Δy_{t-5}	1850	1.014
H_{11}^{α}	1826	1.001	Δy_{t+2}	1767	0.968	Δa_{t+5}	1826	1.001
H_{22}^{α}	1820	0.997	Δa_{t-3}	1858	1.018	Δy_{t+5}	1850	1.014
Δa_t	1823	0.999	Δy_{t-3}	1808	0.991	Δa_{t-6}	1850	1.014
Δy_t	1850	1.014	Δa_{t+3}	1847	1.012	Δy_{t-6}	1826	1.001
Δa_{t-1}	1839	1.008	Δy_{t+3}	1820	0.997	Δa_{t+6}	1839	1.008
Δy_{t-1}	1866	1.022	Δa_{t-4}	1830	1.003	Δy_{t+6}	1866	1.022

Table A.2. Raftery-Lewis Diagnostics for each parameter in the chain. The minimum number of draws under the assumption of i.i.d. draws would be 1825. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

MS-VAR Estimation

In this appendix we provide details on the estimation of the MS-VAR 6. As stated above, we take the regime sequence as given based on our estimates for the breaks in cay^{MS}. Specifically, we choose the particular regime sequence $\hat{\xi}^{\alpha,T} = \{\hat{\xi}_1^{\alpha}, ..., \hat{\xi}_T^{\alpha}\}$ that is most likely to have occurred, given our estimated posterior mode parameter values for $\boldsymbol{\theta}$. This sequence is computed as follows. First, we run Hamilton's filter to get the vector of filtered probabilities $\pi_{t|t}$, $t = 1, 2, ..., T$. The Hamilton filter can be expressed iteratively as

$$\pi_{t|t} = \frac{\pi_{t|t-1} \odot \eta_t}{\mathbf{1}' \left(\pi_{t|t-1} \odot \eta_t \right)}$$
$$\pi_{t+1|t} = \boldsymbol{H}^{\alpha} \pi_{t|t}$$

where η_t is a vector whose j-th element contains the conditional density $p(c_t|\xi_t^{\alpha} = j, x_{M,t}, x_{F,t}; \boldsymbol{\theta})$, the symbol $\odot$ denotes element by element multiplication, and $\mathbf{1}$ is a vector with all elements equal to 1. The final term, $\pi_{T|T}$ is returned with the final step of the filtering algorithm. Then, a recursive algorithm can be implemented to derive the other smoothed probabilities:

$$\pi_{t|T} = \pi_{t|t} \odot \left[\boldsymbol{H}^{\alpha\prime} \left(\pi_{t+1|T} \left(\div \right) \pi_{t+1|t} \right) \right]$$

where $(\div)$ denotes element by element division. To choose the regime sequence most likely to have occurred given our parameter estimates, consider the recursion in the next to last period $t = T - 1$:

$$\pi_{T-1|T} = \pi_{T-1|T-1} \odot \left[\boldsymbol{H}^{\alpha\prime} \left(\pi_{T|T} \left(\div \right) \pi_{T|T-1} \right) \right].$$

We first take $\pi_{T|T}$ from the Hamilton filter and choose the regime that is associated with the largest probability, i.e., if $\pi_{T|T} = (.9, .1)$, where the first element corresponds to the probability of regime 1, we select $\hat{\xi}_T^{\alpha} = 1$, indicating that we are in regime 1 in period T. We now update $\pi_{T|T} = (1, 0)$ and plug into the right-hand-side above along with the estimated filtered probabilities for $\pi_{T-1|T-1}$, $\pi_{T|T-1}$ and estimated transition matrix $\boldsymbol{H}^{\alpha}$ to get $\pi_{T-1|T}$ on the left-hand-side. Now we repeat the same procedure by choosing the regime for $T - 1$ that has the largest probability at $T - 1$, e.g., if $\pi_{T-1|T} = (.2, .8)$ we select $\hat{\xi}_{T-1}^{\alpha} = 2$, indicating that we are in regime 2 in period $T - 1$, we then update to $\pi_{T-1|T} = (0, 1)$, which is used again on the right-hand-side now

$$\pi_{T-2|T} = \pi_{T-2|T-2} \odot \left[\boldsymbol{H}^{\alpha\prime} \left(\pi_{T-1|T} \left(\div \right) \pi_{T-1|T-2} \right) \right].$$

We proceed in this manner until we have a regime sequence $\hat{\xi}^{\alpha,T}$ for the entire sample $t = 1, 2, ..., T$. Two aspects of this procedure are worth noting. First, it fails if the updated probabilities are exactly $(.5, .5)$. Mathematically this is virtually zero. Second, note that this procedure allows us to choose the most likely regime sequence by using the recursive formula above to update the filtered probabilities sequentially from T to time $t = 1$. This allows us to take into account the time dependence in the regime sequence as dictated by the transition probabilities.

Taking regime sequence as given in this way, we need only estimate the transition matrix and the parameters of the MS-VAR across the two regimes. The model is estimated by using Bayesian methods with flat priors on all parameters. As a first step, we group all the observations that belong to the same regime. Conditional on a regime, we have a fixed coefficients VAR. We can then follow standard procedures to make draws for the VAR parameters as follows.

Rewrite the VAR as

$$
\begin{aligned}
\underset{T\times n}{Y} &= \underset{(T\times k)(k\times n)}{X A_{\xi_t}} + \underset{T\times n}{\varepsilon},\ \xi_t = 1, 2 \\
\varepsilon_t &\sim N\left(0, \Sigma_{\xi_t}\right)
\end{aligned}
$$

where $Y = [Z_1, ..., Z_T]'$, the t-th row of X is $X_t = \left[1, Z'_{t-1}, Z'_{t-2}\right]$, $A_{\xi_t} = \left[c_{\xi_t}, A_{1,\xi_t}, A_{2,\xi_t}\right]'$, the t-th row of ε is ε_t, and where $\Sigma_{\xi_t} = V_{\xi_t} V'_{\xi_t}$. If we specify a Normal-Wishart prior for A_{ξ_t} and V_{ξ_t}:

$$
\begin{aligned}
\Sigma_{\xi_t}^{-1} &\sim W\left(S_0^{-1}/v_0, v_0\right) \\
vec\left(A_{\xi_t}|\Sigma_{\xi_t}\right) &\sim N\left(vec\left(B_0\right), \Sigma_{\xi_t} \otimes N_0^{-1}\right)
\end{aligned}
$$

where $E\left(\Sigma_{\xi_t}^{-1}\right) = S_0^{-1}$, the posterior distribution is still in the Normal-Wishart family and is given by

$$
\begin{aligned}
\Sigma_{\xi_t}^{-1} &\sim W\left(S_T^{-1}/v_T, v_T\right) \\
vec\left(A_{\xi_t}|\Sigma_{\xi_t}\right) &\sim N\left(vec\left(B_T\right), \Sigma_{\xi_t} \otimes N_T^{-1}\right)
\end{aligned}
$$

Using the estimated regime sequence $\xi^{\alpha,T}$ we can group all the observations that pertain to the same regime i. Therefore the parameters of the posterior are computed as

$$
\begin{aligned}
v_T &= T_i + v_0, \quad N_T = X'_i X_i + N_0 \\
B_T &= N_T^{-1}\left(N_0 B_0 + X'_i X_i \widehat{B}_{MLE}\right) \\
S_T &= \frac{v_0}{v_T} S_0 + \frac{T_i}{v_T} \widehat{\Sigma}_{MLE} + \frac{1}{v_T}\left(\widehat{B}_{MLE} - \widehat{B}_0\right)' N_0 N_T^{-1} X'_i X_i \left(\widehat{B}_{MLE} - \widehat{B}_0\right) \\
\widehat{B}_{MLE} &= \left(X'_i X_i\right)^{-1}\left(X'_i Y_i\right),\ \widehat{\Sigma}_{MLE} = \frac{1}{T_i}\left(Y_i - X_i \widehat{B}_{MLE}\right)'\left(Y_i - X_i \widehat{B}_{MLE}\right),
\end{aligned}
$$

where T_i, Y_i, X_i denote the number and sample of observations in regime i. We choose flat priors ($v_0 = 0$, $N_0 = 0$) so the expressions above coincide with the MLE estimates using observations in regime i:

$$
v_T = T_i,\ N_T = X'_i X_i,\ B_T = \widehat{B}_{MLE},\ S_T = \widehat{\Sigma}_{MLE}.
$$

Armed with these parameters in each regime, we can make draws from the posterior distributions for $\Sigma_{\xi_t}^{-1}$ and A_{ξ_t} in regime i to characterize parameter uncertainty about these parameters.

Given that we condition the MS-VAR estimates on the most likely regime sequence $\hat{\xi}^{\alpha,T}$ for cay^{MS}, it is still of interest to estimate the elements of the transition probability matrix for the MS-VAR parameters, $\boldsymbol{H}^A$, conditional on this regime sequence. Note that $\boldsymbol{H}^A$ can be different from $\boldsymbol{H}^\alpha$ because the former is based on a particular regime sequence $\hat{\xi}^{\alpha,T}$, while the latter reflects the entire posterior distribution for $\xi^{\alpha,T}$. The estimated transition matrix

$\boldsymbol{H}^A$ can in turn be used to compute expectations taking into account the possibility of regime change (see the next subsection). Because we impose the regime sequence to be the same as that estimated for cay^{MS}, the posterior of $\boldsymbol{H}^A$ only depends on $\xi^{\alpha,T} = \xi$ and does not depend on other parameters of the model. The posterior has a Dirichlet distribution if we assume a prior Dirichlet distribution.[16] For each column of $\boldsymbol{H}^A$ the posterior distribution is given by:

$$\boldsymbol{H}^A(:,i) \sim D(a_{ii} + \eta_{ii,r+1}, a_{ij} + \eta_{ij,r+1})$$

where $\eta_{ij,r+1}$ denotes the number of transitions from regime i to regime j based on $\xi^{\alpha,T}$, while a_{ii} and a_{ij} the corresponding priors. With flat priors, we have $a_{ii} = 0$ and $a_{ij} = 0$. Armed with this posterior distribution, we can characterize uncertainty about $\boldsymbol{H}^A$.

Expectations and Economic Uncertainty

In this appendix we explain how expectations and economic uncertainty are computed for variables in the MS-VAR. More details can be found in Bianchi (2016). Consider the following first-order MS-VAR:

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t}\varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{11}$$

and suppose that we are interested in $\mathbb{E}_0(Z_t) = \mathbb{E}(Z_t|\mathbb{I}_0)$ with $\mathbb{I}_0$ being the information set available at time 0. Note that the first-order VAR is not restrictive because any VAR with $l > 1$ lags can be rewritten as above by using the first-order companion form, and the methods below applied to the companion form.

Let n be the number of variables in the VAR of the previous Appendix section. Let m be the number of Markov-switching states. Define the $mn \times 1$ column vector q_t as:

$$\underset{mn\times 1}{q_t} = \left[q_t^{1\prime}, ..., q_t^{m\prime}\right]'$$

where the individual $n \times 1$ vectors $q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t=i}\right) \equiv \mathbb{E}\left(Z_t 1_{\xi_t=i}|\mathbb{I}_0\right)$ and $1_{\xi_t=i}$ is an indicator variable that is one when regime i is in place and zero otherwise. Note that:

$$q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t=i}\right) = \mathbb{E}_0\left(Z_t|\xi_t = i\right)\pi_t^i$$

where $\pi_t^i = P_0(\xi_t = i) = P(\xi_t = i|\mathbb{I}_0)$. Therefore we can express $\mu_t = \mathbb{E}_0(Z_t)$ as:

$$\mu_t = \mathbb{E}_0(Z_t) = \textstyle\sum_{i=1}^m q_t^i = wq_t$$

where the matrix $\underset{n\times mn}{w} = [I_n, ..., I_n]$ is obtained placing side by side m n-dimensional identity matrices. Then the following proposition holds:

[16] The Dirichlet distribution is a generalization of the beta distribution that allows one to potentially consider more than 2 regimes. See e.g., Sims and Zha (2006).

Proposition 1 *Consider a Markov-switching model whose law of motion can be described by (11) and define* $q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t=i}\right)$ *for* $i = 1...m$. *Then* $q_t^j = c_j \pi_t^j + \sum_{i=1}^m A_j q_{t-1}^i h_{ji}$.

It is then straightforward to compute expectations conditional on the information available at a particular point in time. Suppose we are interested in $\mu_{t+s|t} \equiv \mathbb{E}_t\left(Z_{t+s}\right)$, i.e. the expected value for the vector Z_{t+s} conditional on the information set available at time t. If we define:

$$q_{t+s|t} = \left[q_{t+s|t}^{1\prime}, ..., q_{t+s|t}^{m\prime}\right]'$$

where $q_{t+s|t}^i = \mathbb{E}_t\left(Z_{t+s} 1_{\xi_{t+s}=i}\right) = \mathbb{E}_t\left(Z_{t+s}|\xi_{t+s} = i\right) \pi_{t+s|t}^i$, where $\pi_{t+s|t}^i \equiv P\left(\xi_{t+s} = i|\mathbb{I}_t\right)$, we have

$$\mu_{t+s|t} = \mathbb{E}_t\left(Z_{t+s}\right) = w q_{t+s|t}, \tag{12}$$

where for $s \geq 1$, $q_{t+s|t}$ evolves as:

$$q_{t+s|t} = C\pi_{t+s|t} + \Omega q_{t+s-1|t} \tag{13}$$

$$\pi_{t+s|t} = H\pi_{t+s-1|t} \tag{14}$$

with $\pi_{t+s|t} = \left[\pi_{t+s|t}^1, ..., \pi_{t+s|t}^m\right]'$, $\Omega = bdiag\left(A_1, ..., A_m\right)\left(H \otimes I_n\right)$, and $\underset{mn\times m}{C} = bdiag\left(c_1, ..., c_m\right)$, where e.g., c_1 is the $n \times 1$ vector of constants in regime 1, $\otimes$ represents the Kronecker product and $bdiag$ is a matrix operator that takes a sequence of matrices and use them to construct a block diagonal matrix.

Similar formulas hold for the second moments. Before proceeding, let us define the vectorization operator $\varphi\left(X\right)$ that takes the matrix X as an input and returns a column vector stacking the columns of the matrix X on top of one another. We will also make use of the following result: $\varphi\left(X_1X_2X_3\right) = \left(X_3' \otimes X_1\right)\varphi\left(X_2\right)$.

Define the vector $n^2m \times 1$ column vector Q_t as:

$$Q_t = \left[Q_t^{1\prime}, ..., Q_t^{m\prime}\right]'$$

where the $n^2 \times 1$ vector Q_t^i is given by $Q_t^i = \varphi\left[\mathbb{E}_0\left(Z_tZ_t'1_{\xi_t=i}\right)\right]$. This implies that we can compute the vectorized matrix of second moments $M_t = \varphi\left[\mathbb{E}_0\left(Z_tZ_t'\right)\right]$ as:

$$M_t = \varphi\left[\mathbb{E}_0\left(Z_tZ_t'\right)\right] = \sum_{i=1}^m Q_t^i = WQ_t$$

where the matrix $W = \left[I_{n^2}, ..., I_{n^2}\right]$ is obtained placing side by side m n^2-dimensional identity matrices. We can then state the following proposition:

Proposition 2 *Consider a Markov-switching model whose law of motion can be described by (11) and define* $Q_t^i = \varphi\left[\mathbb{E}_0\left(Z_tZ_t'1_{\xi_t=i}\right)\right]$, $q_t^i = \mathbb{E}_0\left[Z_t1_{\xi_t=i}\right]$, *and* $\pi_t^i = P_0\left(\xi_t = i\right)$, *for* $i = 1...m$. *Then* $Q_t^j = \left[\widehat{cc}_j + \widehat{VV}_j\varphi\left[I_k\right]\right]\pi_t^j + \sum_{i=1}^m\left[\widehat{AA}_jQ_{t-1}^i + \widehat{DAC}_jq_{t-1}^i\right]h_{ji}$, *where* $\widehat{cc}_j = \left(c_j \otimes c_j\right)$, $\widehat{VV}_j = \left(V_j \otimes V_j\right)$, $\widehat{AA}_j = \left(A_j \otimes A_j\right)$, *and* $\widehat{DAC}_j = \left(A_j \otimes c_j\right) + \left(c_j \otimes A_j\right)$.

It is then straightforward to compute the evolution of second moments conditional on the information available at a particular point in time. Suppose we are interested in $\mathbb{E}_t\left(Z_{t+s}Z'_{t+s}\right)$, i.e. the second moment of the vector Z_{t+s} conditional on the information available at time t. If we define:

$$Q_{t+s|t} = \left[Q^{1\prime}_{t+s|t}, ..., Q^{m\prime}_{t+s|t}\right]'$$

where $Q^i_{t+s|t} = \varphi\left(\mathbb{E}_t\left(Z_{t+s}Z'_{t+s}1_{\xi_{t+s}=i}\right)\right) = \varphi\left(\mathbb{E}_t\left(Z_{t+s}Z'_{t+s}|\xi_{t+s}=i\right)\right)\pi^i_{t+s|t}$, we obtain $\varphi\left(\mathbb{E}_t\left(Z_{t+s}Z'_{t+s}\right)\right) = WQ_{t+s|t}$.Using matrix algebra we obtain:

$$Q_{t+s|t} = \Xi Q_{t+s-1|t} + \widehat{DAC}q_{t+s-1|t} + \widehat{V}c\pi_{t+s|t} \tag{15}$$

$$q_{t+s|t} = C\pi_{t+s|t} + \Omega q_{t+s-1|t},\ \pi_{t+s|t} = H\pi_{t+s-1|t}. \tag{16}$$

where

$$\Xi = bdiag(\widehat{AA}_1, ..., \widehat{AA}_m)(H \otimes I_{n^2}),\ \widehat{V}c = \left[\widehat{VV} + \widehat{cc}\right],\ \widehat{cc} = bdiag(\widehat{cc}_1, ..., \widehat{cc}_m),$$

$$\widehat{VV} = bdiag(\widehat{VV}_1\varphi\left[I_k\right], ..., \widehat{VV}_m\varphi\left[I_k\right]),\ \widehat{DAC} = bdiag(\widehat{DAC}_1, ..., \widehat{DAC}_m)(H \otimes I_n).$$

With the first and second moments at hand, it is then possible to compute the variance s periods ahead conditional on the information available at time t:

$$\varphi\left[\mathbb{V}_t\left(Z_{t+s}\right)\right] = M_{t+s|t} - \varphi\left[\mu_{t+s|t}\mu'_{t+s|t}\right], \tag{17}$$

where $M_{t+s|t} = \varphi\left(\mathbb{E}_t\left(Z_{t+s}Z'_{t+s}\right)\right) = \sum_{i=1}^m Q^i_{t+s|t} = WQ_{t+s|t}$.

To report estimates of (12) and (17) we proceed as follows. Note that $\mu_{t+s|t} = \mathbb{E}_t\left(Z_{t+s}\right) = wq_{t+s|t}$ and $M_{t+s|t}$ depend only on $q_{t+s|t}$ and $Q_{t+s|t}$. Furthermore we can express (13)-(14) and (15)-(16) in a compact form as

$$\widetilde{Q}_{t+s|t} = \widetilde{\Xi}^s\widetilde{Q}_{t|t} \text{ where } \widetilde{\Xi} = \left[\begin{array}{c|cc} \Xi & \widehat{DAC} & \widehat{V}cH \\ \hline & \Omega & CH \\ & & H \end{array}\right], \tag{18}$$

where $\widetilde{Q}_{t+s|t} = \left[Q'_{t+s|t}, q'_{t+s|t}, \pi'_{t+s|t}\right]'$. Armed with starting values $\widetilde{Q}_{t|t} = \left[Q'_{t|t}, q'_{t|t}, \pi'_{t|t}\right]'$ we can then compute (12) and (17) using (18). To obtain $\pi'_{t|t}$ recall that we assume that $\mathbb{I}_t$ includes knowledge of the regime in place at time t, the data up to time t, Z^t, and the VAR parameters for each regime. Given that we assume knowledge of the current regime, $\pi^i_{t|t} \equiv P\left(\xi_t = i|\mathbb{I}_t\right)$ can only assume two values, 0 or 1. As a result $\pi'_{t|t}$ will be $(1,0)$ or $(0,1)$. As a result, and given $Z_t \in \mathbb{I}_t$, $q'_{t|t} = \left[q^{1\prime}_{t|t}, q^{2\prime}_{t|t}\right]'$ with $q^i_{t|t} \equiv \mathbb{E}_t\left(Z_t|\xi_t = i\right)\pi^i_{t|t}$, will be $[Z'_t \cdot 1, Z'_t \cdot 0]'$ or $[Z'_t \cdot 0, Z'_t \cdot 1]'$. Analogously, $Q'_{t|t} = \left[Q^{1\prime}_{t|t}, Q^{2\prime}_{t|t}\right]'$ with $Q^i_{t|t} \equiv \varphi\left(\mathbb{E}_t\left(Z_tZ'_t|\xi_t = i\right)\right)\pi^i_{t|t|}$ will be $\left[\varphi\left(Z_tZ'_t \cdot 1\right)', \varphi\left(Z_tZ'_t \cdot 0\right)'\right]'$ or $\left[\varphi\left(Z_tZ'_t \cdot 0\right)', \varphi\left(Z_tZ'_t \cdot 1\right)'\right]'$.

6.2 Conditional Steady State

Consider a MS-VAR:

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t} \varepsilon_t$$

where Z_t is a column vector containing n variables observable at time t and $\xi_t = 1, ..., m$, with m the number of regimes, evolves following the transition matrix H. If the MS-VAR has more than one lag, the companion form can be used to recast the model as illustrated above.

The conditional steady state corresponds to the expected value for the vector Z_t conditional on being in a particular regime. This is computed by imposing that a certain regime is in place forever:

$$\mathbb{E}_i (Z_t) = \overline{\mu}_i = (I_n - A_i)^{-1} c_i \tag{19}$$

where I_n is an identity matrix with the appropriate size. Note that unless the VAR coefficients imply very slow moving dynamics, after a switch from regime j to regime i, the variables of the VAR will converge (in expectation) to $\mathbb{E}_i (Z_t)$ over a finite horizon. If there are no further switches, we can then expect the variables to fluctuate around $\mathbb{E}_i (Z_t)$. Therefore, the conditional steady states can also be thought as the values to which the variables converge if regime i is in place for a long enough period of time.

6.3 Book-to-Market Ratio

We use the methods and assumptions of the previous subsection to obtain the present value decomposition of the book to market ratio. Consider an MS-VAR:

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t} \varepsilon_t$$

where Z_t is a column vector containing n variables observable at time t and $\xi_t = 1, ..., m$, with m the number of regimes, evolves following the transition matrix H. If the MS-VAR has more than one lag, the companion form can be used to recast the model as illustrated above.

Define the column vectors q_t and π_t:

$$q_t = \left[q_t^{1'}, ..., q_t^{m\prime} \right]', q_t^i = \mathbb{E}_0 \left(Z_t 1_{\xi_t = i} \right), \pi_t = \left[\pi_t^1, ..., \pi_t^m \right]',$$

where $\pi_t^i = P_0 (\xi_t = i)$ and $1_{\xi_t = i}$ is an indicator variable that is equal to 1 when regime i is in place and zero otherwise. The law of motion for $\widetilde{q}_t = [q_t', \pi_t']'$ is then given by

$$\underbrace{\begin{bmatrix} q_t \\ \pi_t \end{bmatrix}}_{\widetilde{q}_t} = \underbrace{\begin{bmatrix} \Omega & CH \\ & H \end{bmatrix}}_{\widetilde{\Omega}} \begin{bmatrix} q_{t-1} \\ \pi_{t-1} \end{bmatrix}$$

where $\pi_t = [\pi_{1,t}, ..., \pi_{m,t}]'$, $\Omega = bdiag(A_1, ..., A_m) H$, and $C = bdiag(c_1, ..., c_m)$. Recall that:

$$\mathbb{E}_0(Z_t) = \sum_{i=1}^{m} q_t^i = wq_t, \;\; w = \left[\underbrace{I_n, ..., I_n}_{m}\right]$$

To compute the present value decomposition of the book-to-market ratio, define:

$$\begin{aligned} q^i_{t+s|t} &= \mathbb{E}_t\left(Z_{t+s} 1_{\xi_{t+s}=i}\right) = \mathbb{E}\left(Z_{t+s} 1_{\xi_{t+s}=i} | \mathbb{I}_t\right) \\ 1'_x &= [0, ...1, ...0, 0, 0]', \;\; mn = m * n \end{aligned}$$

where $\mathbb{I}_t$ contains all the information that agents have at time t, including the probability of being in one of the m regimes. Note that $q^i_{t|t} = Z_t \pi^i_t$.

Now consider the formula from Vuolteenaho (1999):

$$\theta_t = \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t r_{t+1+j} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t f_{t+1+j} - \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t e^*_{t+1+j}$$

Given that our goal is to assess if assets with different risk profiles are affected differently by the breaks in the long-term interest rates, we are going to focus on the *difference* between the book-to-market ratios. Specifically, given two portfolios x and y, we are interested in how the difference in their book-to-market ratios, $\theta_{x,t} - \theta_{y,t}$, varies across the two regimes:

$$\underbrace{\theta_{x,t} - \theta_{y,t}}_{\text{Spread in BM ratios}} = \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(r_{x,t+1+j} - r_{y,t+1+j}\right)}_{\text{PDV of the difference in expected excess returns}} - \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(e^*_{x,t+1+j} - e^*_{y,t+1+j}\right)}_{\text{PDV of the difference in expected earnings}}$$

If then we want to correct the spread in BM ratios by taking into account expected earnings, we have:

$$\underbrace{\theta_{x,t} - \theta_{y,t} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(e^*_{x,t+1+j} - e^*_{y,t+1+j}\right)}_{\text{Spread in BM ratios corrected for earnings}} = \underbrace{\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(r_{x,t+1+j} - r_{y,t+1+j}\right)}_{\text{PDV of the expected difference in excess returns}} \tag{20}$$

Suppose that we have estimated a MS-VAR that includes the difference in excess returns, $r_{xy,t} \equiv r_{x,t} - r_{y,t}$. Then the right hand side of (20) can be computed as:

$$\begin{aligned} \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(r_{xy,t+1+j}\right) &= \sum_{j=0}^{\infty} \rho^j 1'_{r_{xy}} w q_{t+1+j|t} \\ &= 1'_{r_{xy}} w \left(I - \rho\Omega\right)^{-1} \left[\Omega q_{t|t} + C\left(I - \rho H\right)^{-1} H \pi_{t|t}\right]. \end{aligned}$$

Therefore, we have:

$$\widetilde{\theta}_{xy,t} \equiv \underbrace{\widetilde{\theta}_{x,t} - \widetilde{\theta}_{y,t} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t\left(e^*_{x,t+1+j} - e^*_{y,t+1+j}\right)}_{\text{Spread in BM ratios corrected for earnings}} = 1'_{r_{xy}} w \left(I - \rho\Omega\right)^{-1} \left[\Omega q_{t|t} + C\left(I - \rho H\right)^{-1} H \pi_{t|t}\right]$$

where we have used $\widetilde{\theta}_{xy,t}$ to define the spread in BM ratios corrected for earnings.

Note that we can also compute long term values, both conditioning on a regime or not, because the present value decomposition is just a function of the state vector and the regime probabilities. Therefore:

$$\overline{\overline{\theta}}_{xy,i} = 1'_{r_{xy}} w \left(I - \rho\Omega\right)^{-1} \left[\Omega \overline{q}_i + C \left(I - \rho H\right)^{-1} H \overline{\pi}_i\right]$$

where $\overline{\pi}_i = 1_i$ and $\overline{q}_i$ is the conditional steady state of the VAR vector Z_t for regime i. This corresponds to the values to which the variables converge if regime i is in place for a prolonged period of time (see below). Note that $\overline{\overline{\theta}}_{xy,i}$ is computed by conditioning to *initially* being in regime i, but taking into account that there might be regime changes in the future.

Variable	NSE	RNE	Variable	NSE	RNE	Variable	NSE	RNE
α_1	0.000131	1	Δa_{t+1}	0.000263	1	Δy_{t-4}	0.000526	1
α_2	0.000131	1	Δy_{t+1}	0.00053	1	Δa_{t+4}	0.000256	1
β_a	0.000074	1	Δa_{t-2}	0.000261	1	Δy_{t+4}	0.000521	1
β_y	0.000085	1	Δy_{t-2}	0.000572	1	Δa_{t-5}	0.000264	1
σ	0	1	Δa_{t+2}	0.000258	1	Δy_{t-5}	0.000524	1
H_{11}^{α}	0.000069	1	Δy_{t+2}	0.000547	1	Δa_{t+5}	0.000252	1
H_{22}^{α}	0.000053	1	Δa_{t-3}	0.000278	1	Δy_{t+5}	0.000534	1
Δa_t	0.000263	1	Δy_{t-3}	0.000632	1	Δa_{t-6}	0.000275	1
Δy_t	0.000529	1	Δa_{t+3}	0.000255	1	Δy_{t-6}	0.000518	1
Δa_{t-1}	0.000252	1	Δy_{t+3}	0.000537	1	Δa_{t+6}	0.000238	1
Δy_{t-1}	0.000521	1	Δa_{t-4}	0.000259	1	Δy_{t+6}	0.000525	1

Table A.3 The table reports the numerical standard error (NSE) and the relative numerical efficiency (RNE) computed based on Geweke (1992). Values for NSE close to zero and values for RSE close to 1 are indicative of convergence. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

References

Acharya, V. V., and H. Naqvi (2016): "On reaching for yield and the coexistence of bubbles and negative bubbles," *Available at SSRN 2618973*.

Andrews, D. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Barro, R. J., and A. Mollerus (2014): "Safe assets," National Bureau of Economic Research No. w20652.

Bernanke, B. S., and K. N. Kuttner (2005): "What Explains the Stock Market's Reaction to Federal Reserve Policy?," *Journal of Finance*.

Bianchi, F. (2013): "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," *The Review of Economic Studies*, 80(2), 463–490.

——— (2016): "Methods for measuring expectations and uncertainty in Markov-switching models," *Journal of Econometrics*, 190(1), 79–99.

Bianchi, F., and C. Ilut (2015): "Monetary/Fiscal Policy Mix and Agents' Beliefs," CEPR discussion paper 9645, NBER working paper 20194.

Caballero, R. J., and E. Farhi (2014): "The Safety Trap," Discussion paper, National Bureau of Economic Research No. w19927.

Campbell, J. Y., C. Pflueger, and L. M. Viceira (2014): "Monetary policy drivers of bond and equity risks," Discussion paper, National Bureau of Economic Research.

Campbell, J. Y., and S. Thompson (2008): "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, 21, 1509–1531.

Choi, J., and M. Kronlund (2015): "Reaching for yield or playing it safe? Risk taking by bond mutual funds," Available at SSRN: https://ssrn.com/abstract=2669465.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen (2015): "Stock returns over the FOMC cycle," *Available at SSRN 2687614*.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 147–180.

Cook, T. a. H. (1989): "The effect of changes in the federal funds rate target on market interest rates in the 1970s," *Journal of Monetary Economics*, 24(3), 331–351.

Daniel, K. D., and T. J. Moskowitz (2013): "Momentum Crashes," Swiss Finance Institute Research Paper.

Di Maggio, M., and M. T. Kacperczyk (2015): "The unintended consequences of the zero lower bound policy," Columbia Business School Research Paper.

Drechsler, I., A. Savov, and P. Schnabl (2014): "A model of monetary policy and risk premia," National Bureau of Economic Research No. w20141.

Fama, E. F., and K. R. French (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116(1), 1–22.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003): "Spurious regressions in financial economics?," *Journal of Finance*, 58(4), 1393–1413.

Galí, J. (2014): "Monetary Policy and Rational Asset Price Bubb," *The American Economic Review*, 104(3), 721–752.

Geweke, J. F. (1992): "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," *Bayesian Statistics*, 4, 169–193.

Goyal, A., and I. Welch (2003): "Predicting the Equity Premium with Dividend Ratios," *Management Science*, 49(5), 639–654.

——— (2008): "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21(4), 1455–1508.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005): "The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models," *The American Economic Review*, 95(1), 425–436.

Hall, R. E. (2016): "Understanding the Decline in the Safe Real Interest Rate," National Bureau of Economic Research No. w22196.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, NY.

Han, H.-L., and M. Ogaki (1997): "Consumption, Income and Cointegration," *International Review of Economics adn Finance*, 6(2), 107–117.

Hanson, S. G., and J. C. Stein (2015): "Monetary policy and long-term real rates," *Journal of Financial Economics*, 115(3), 429–448.

Harrison, J. M., and D. M. Kreps (1978): "Speculative investor behavior in a stock market with heterogeneous expectations," *The Quarterly Journal of Economics*, pp. 323–336.

KIM, C.-J., AND C. R. NELSON (1999): *State-Space Models with Regime Switching.* The MIT Press, Cambridge, Massachusetts.

KOIJEN, R. S. J., AND S. VAN NIEUWERBURGH (2011): "Predictability of Returns and Cash Flows," *Annual Review of Financial Economics*, 3, 467–491.

LETTAU, M., AND S. C. LUDVIGSON (2001): "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance*, 56(3), 815–849.

——— (2004): "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94(1), 276–299.

——— (2010): "Measuring and Modeling Variation in the Risk-Return Tradeoff," in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia, and L. P. Hansen, vol. 1, pp. 617–90. Elsevier Science B.V., North Holland, Amsterdam.

LETTAU, M., AND S. C. LUDVIGSON (2013): "Shocks and Crashes," in *National Bureau of Economics Research Macroeconomics Annual: 2013*, ed. by J. Parker, and M. Woodford, vol. 28, pp. 293–354. MIT Press, Cambridge and London.

LETTAU, M., S. C. LUDVIGSON, AND J. A. WACHTER (2008): "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," *Review of Financial Studies*, 21(4), 1653–1687.

LETTAU, M., AND S. VAN NIEUWERBURGH (2008): "Reconciling the Return Predictability Evidence: In-Sample Forecasts, Out-of-Sample Forecasts, and Parameter Instability," *Review of Financial Studies*, 21(4), 1607–1652.

LEWELLEN, J. W. (2004): "Predicting Returns With Financial Ratios," *Journal of Financial Economics*, 74, 209–235.

LUBIK, T. A., AND F. SCHORFHEIDE (2004): "Testing for indeterminacy: an application to US monetary policy," *American Economic Review*, pp. 190–217.

LUCCA, D. O., AND E. MOENCH (2015): "The Pre-FOMC Announcement Drift," *The Journal of Finance*, 70(1), 329–371.

MCCONNELL, M. M., AND G. PEREZ-QUIROS (2000): "Output Fluctuations in the United States: What Has Changed Since the early 1980s?," *American Economic Review*, 90(5), 1464–1476.

MULLER, U. K., AND M. W. WATSON (2008): "Testing Models of Low Frequency Variability," *Econometrica*, 76, 155–74.

——— (2013): "Measuring Uncertainty About Long-Run Predictions," Unpublished paper, Princeton University.

Neuhierl, A., and M. Weber (2016): "Monetary policy and the stock market: Time-series evidence," *Available at SSRN*.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Ogaki, M., and J. Y. Park (1997): "A Cointegration Approach to Estimating Preference Parameters," *Journal of Econometrics*, 82(1), 107–134.

Park, J. Y. (1990): "Testing for Unit Roots and Cointegration by Variable Addition," *Advances in Econometrics*, 8, 107–133.

——— (1992): "Canonical Cointegrating Regressions," *Econometrica*, 60(1), 119–143.

Park, J. Y., and M. Ogaki (1991): "Inference in Cointegrated Models using VAR prewhitening to estimate Shortrun Dynamics," Rochester Center for Economic Research Working Paper No. 281, University of Rochester.

Phillips, P., and S. Ouliaris (1990): "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58, 165–193.

Raftery, A., and S. Lewis (1992): "How many iterations in the Gibbs sampler?," *Bayesian Statistics*, 4, 763–773.

Rajan, R. G. (2006): "Has finance made the world riskier?," *European Financial Management*, 12(4), 499–533.

——— (2013): "A step in the dark: unconventional monetary policy after the crisis," *Andrew Crockett Memorial Lecture, BIS, Basel*, 23.

Scheinkman, J. A., and W. Xiong (2003): "Overconfidence and speculative bubbles," *Journal of political Economy*, 111(6), 1183–1220.

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in US Monetary Policy?," *American Economic Review*, 91(1), 54–81.

Stock, J. H., and M. W. Watson (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.

——— (2002): "Has the Business Cycle Change and Why?," in *NBER Macroeconomics Annual: 2002*, ed. by M. Gertler, and K. Rogoff. MIT Press, Cambridge, MA.

Taylor, J. B. (2007): "Housing and Monetary Policy," in *Housing, Housing Finance, and Monetary Policy: Proceedings of the Federal Reserve Bank of Kansas City's Symposium, held at Jackson Hole, Wyoming, August 30-Sept 1, 2007*, pp. 463–476. Federal Reserve Bank of Kansas City.

van Binsbergen, J. H., and R. S. Koijen (2010): "Predictive regressions: A present-value approach," *The Journal of Finance*, 65(4), 1439–1471.

Vuolteenaho, T. (2000): "Understanding the Aggregate Book-Market Ratio and its Implications to Current Equity-Premium Expectations," Unpublished paper, Harvard University.

Watson, M. W. (2013): "Comment on "Shocks and Crashes"," in *National Bureau of Economics Macroeconomics Annual: 2013*, ed. by J. Parker, and M. Woodford, pp. 367–378. MIT Press, Cambridge and London.

Tables and Figures

	Mode	Mean	5%	95%
α_1	0.9186	0.9153	0.8853	0.9460
α_2	0.8808	0.8767	0.8467	0.9077
$\alpha_1 - \alpha_2$	0.0378	0.0385	0.0358	0.0413
β_a	0.2606	0.2679	0.2505	0.2852
β_y	0.6156	0.6071	0.5873	0.6270
σ	0.0080	0.0087	0.0080	0.0094
H_{11}^{α}	0.9900	0.9901	0.9705	0.9995
H_{22}^{α}	0.9925	0.9923	0.9771	0.9996

Table 1: Posterior modes, means, and 90% error bands of the parameters of the Markov-switching cointegrating relation. Flat priors are used on all parameters of the model. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Parameter Estimates: cay^{FC}

α	γ_a	γ_y
0.8706 (0.0345)	0.1246 (0.0150)	0.7815 (0.0168)

Table 2: Parameter estimates for the fixed coefficient cointegrating relation. Standard errors are in parantheses. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Cointegration Tests

		Dickey–Fuller t-statistic				*Critical values*	
	$Persistence\ cay$	$Lag = 1$	$Lag = 2$	$Lag = 3$	$Lag = 4$	5%	10%
MS	0.8131	-4.7609	-4.4168	-4.4586	-4.7618	-3.80	-3.52
FC	0.9377	-2.2911	-2.1556	-1.8894	-1.6583	-3.80	-3.52

Table 3: The first column reports the first-order autoregressive coefficient obtained regressing cay_t on its own lagged value and a constant. The next four columns report augmented Dickey-Fuller t-statistics $(\hat{\rho} - 1)/\hat{\sigma}_{\hat{\rho}}$, where $\hat{\rho}$ is the estimated value for the autoregressive coeffcient used to test the null hypothesis of no cointegration. This test is applied to estimates of the cointegrating residual, cay_t. We include up to four lags of the first difference of cay_t. The critical values for the test when applied to cointegrating residual are reported in the last two columns and are taken from Phillips and Ouliaris (1990). The results for cay_t^{MS} do not account for sampling error in the estimated Markov-switching mean. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Canonical Cointegrating Regression Results

$\widehat{\gamma}_a$	$\widehat{\gamma}_y$	$H(0,1)$
0.0774	0.8690	0.5720
(0.0665)	(0.0731)	(0.4495)

Table 4: Test results for the null of cointegration for standard, fixed-coefficient *cay*. A rejection of the null at the 5 percent level is warranted if the p-value for the $H(0,1)$ statistic is less than 0.05. Ogaki and Park's (1991) VAR pre-whitening method with Andrews' (1991) automatic bandwidth parameter estimator was used to estimate long-run covariance parameters. The parameters $\widehat{\gamma}_a$ and $\widehat{\gamma}_y$ are estimated cointegrating parameters on a and y, respectively. Standard errors are in parentheses. $H(0,1)$ has a $\chi^2(1)$ distribution. p values for this statistic are in parentheses. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Long Horizon Forecasting Regressions: Stock Returns

h-period regression: $\sum_{i=1}^{h}(r_{t+i} - r_{f,t+i}) = k + \gamma \ \ z_t + \epsilon_{t,t+h}$

$z_t =$	Horizon h (in quarters) 1	4	8	12	16
	Full sample				
cay^{FC}	**0.60**	**2.26**	**4.16**	**5.68**	**7.42**
	(2.00)	(2.21)	(2.47)	(2.73)	(3.71)
	[0.01]	[0.05]	[0.10]	[0.14]	[0.20]
cay^{MSfilt}	**1.54**	**6.38**	**11.60**	**13.56**	**13.61**
	(4.07)	(5.22)	(6.53)	(6.03)	(6.18)
	[0.04]	[0.18]	[0.35]	[0.37]	[0.34]
cay^{MS}	**1.49**	**6.83**	**11.88**	**13.79**	**13.78**
	(3.86)	(6.08)	(6.63)	(6.11)	(6.25)
	[0.04]	[0.21]	[0.36]	[0.38]	[0.34]
	Sub-sample 1981Q1-2013Q3, recursive				
cay^{FC}	0.17	1.00	2.48	3.96	6.39
	(0.48)	(0.83)	(1.04)	(1.18)	(1.82)
	[-0.01]	[0.00]	[0.03]	[0.06]	[0.11]
cay^{FCrec}	0.30	1.67	**4.04**	**6.16**	**8.10**
	(0.97)	(1.65)	(2.29)	(2.79)	(4.17)
	[0.00]	[0.04]	[0.16]	[0.27]	[0.41]
cay^{MSrec}	0.41	2.13	**6.01**	**8.65**	**10.33**
	(1.10)	(1.92)	(2.73)	(3.51)	(5.17)
	[0.00]	[0.04]	[0.21]	[0.31]	[0.37]

Table 5: This tables reports the results from regressions of of h-period-ahead CRSP-VW returns in excess of a 3-month Treasury-bill rate, $r_{f,t}$, on the variable listed in the first column. cay^{FC} is the fixed-coefficient consumption-wealth ratio; cay^{MSfilt} denotes the Markov-switching version of *cay* using filtered probabilities and cay^{MS} denotes the benchmark Markov-switching *cay* using smoothed probabilities. The bottom panel reports results from regressions using recursively estimated versions of *cay*, in which all parameters are estimated using data up to time t rather than using the full sample. The models are first estimated on data from 1952Q1-1970Q1. We then recursively add observations and reestimate the *cay* variables over expanding sub-samples using data only up to the end of that subsample, continuing in this way until the end of the sample, 2013:Q3. Results are reported for the subsample since 1980. cay^{FCrec} denotes the fixed coefficient *cay* estimated recursively, while cay^{MSrec} denotes the Markov-switching *cay* estimated recursively using smoothed probabilities. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses), and adjusted R^2 statistics in square brackets. Significant coefficients based on a t-test at the 5% significance level are highlighted in bold face. The full sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Out-Of-Sample Forecasts

h-period regression: $\sum_{i=1}^{h}(r_{t+i} - r_{f,t+i}) = k + \gamma \; z_t + \epsilon_{t,t+h}$

	Horizon h (in quarters)				
$z_t =$	1	4	8	12	16
	Mean-squared errors				
const	0.75	3.08	5.48	7.92	9.73
$r - r_f$	0.71	2.99	5.32	7.67	9.36
cay^{FC}	0.71	2.90	4.67	6.74	7.36
cay^{MSfilt}	0.70	2.47	2.64	3.01	3.72
cay^{MS}	0.70	2.35	2.53	2.92	3.68
cay^{FCrec}	0.72	2.87	4.38	5.72	6.61
cay^{MSrec}	0.71	2.86	4.49	5.75	6.14

Table 6: This tables reports the mean-squared forecast errors from out-of-sample h-period-ahead forecasts of CRSP-VW returns in excess of a 3-month Treasury-bill rate using 60-quarter rolling subsamples. The single predictor variable in each regression is listed in the first column. The forecasting regression is first estimated on data from 1952Q1-1980Q1, and forecasts are made over the next h periods. We then repeat this forecasting regression using data from the next 60 quarters of the sample, continuing in this way until the end of the sample, 2013:Q3. Mean-square-errors are reported for the subsample since 1980. cay^{FC} is the fixed-coefficient consumption-wealth ratio, cay^{MSfilt} and cay^{MS} are the Markov-switching *cay* variables using filtered and smoothed probabilities, respectively, cay^{FCrec} is the recursively estimated *cay* with fixed coefficients, and cay^{MSrec} is the recursively estimated Markov-switching *cay*. The recursive estimates use data only up to time t. The full sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Summary statistics for the Real Interest Rate and GDP growth

	Real Interest Rate		GDP growth	
	Regime 1	Regime 2	Regime 1	Regime 2
Conditional Means	3.5364 (3.5093,3.5631)	0.5932 (0.5745,0.6112)	3.5085 (3.4838,3.5313)	2.9154 (2.8950,2.9363)
Conditional St. Deviations	2.0921 (1.9629,2.2404)	1.5920 (1.5056,1.6863)	1.9693 (1.8642,2.0719)	2.6450 (2.5025,2.8112)

Table 7: This table reports the mean and standard deviation for the real interest rate and GDP growth based on the VAR estimates conditional on staying in each regime. In parentheses we report the 65% posterior credible sets. The sample spans the period 1955:Q3-2013:Q3.

Annualized Sharpe Ratios and Mean Returns

Portfolio	SR	Mean	Portfolio	SR	Mean	Portfolio	SR	Mean
V-G (S1)	0.6225	0.1047	V-RF (S1)	0.5271	0.1346	G-RF (S1)	0.0909	0.0299
V-G (S2)	0.3807	0.0637	V-RF (S2)	0.5072	0.1202	G-RF (S2)	0.1958	0.0565
V-G (S3)	0.4025	0.0671	V-RF (S3)	0.5786	0.1260	G-RF (S3)	0.2255	0.0589
V-G (S4)	0.1681	0.0271	V-RF (S4)	0.4312	0.0974	G-RF (S4)	0.2973	0.0703
V-G (S5)	0.1899	0.0285	V-RF (S5)	0.3978	0.0785	G-RF (S5)	0.2733	0.0500
W-L	0.6446	0.1588	W-RF	0.5689	0.1315	L-RF	−0.0859	−0.0273

Table 8: The table reports annualized Sharpe ratios, "SR," and mean returns, "Mean," for different portfolios. The Sharpe ratio is defined to be the unconditional mean return divided by the standard deviation of the portfolio return. The long-short portfolios "V-G" are the value-growth portfolios in a given size quintile, S1=smallest, S5=largest. long-short portfolios "W-L" are the winner-loser portfolio. For each size category, the return of the V-G portfolio portfolio return is the difference between the return on the extreme value (highest BM ratio) and the return of the extreme growth portfolio (lowest BM ratio). The return of the W-L portfolio return is the difference in returns between the extreme winner (M10) and the extreme loser (M1). The rows denoted "V-RF", "G-RF", "W-RF" and "L-RF" report the same statistics for the value, growth, winner and loser portfolios, respectively, in excess of the risk-free rate. All returns are computed at quarterly frequencies but the Sharpe ratios and mean returns are reported in annualized units. The sample spans the period 1964:Q1-2013:Q3.

Breaks in Book-Market Ratio Spreads

	Size 1 (Small)		Size 2		Size 3	
	Val-Gr	W-L	Val-Gr	W-L	Val-Gr	W-L
Regime 1	2.5402 (2.4791,2.6127)	4.4258 (3.8221,4.8132)	1.5740 (1.4708,1.6530)	4.4525 (3.8418,4.8380)	1.6086 (1.5560,1.6491)	4.3910 (3.8142,4.7666)
Regime 2	2.3591 (2.2143,2.4869)	2.9713 (2.2002,3.7477)	1.1795 (0.9913,1.3387)	2.9560 (2.2761,3.7596)	1.4296 (1.3055,1.5265)	2.9494 (2.1859,3.7089)
Diff-in-Diff	0.1832 (0.0489,0.3441)	1.3134 (0.8419,1.9107)	0.3817 (0.2537,0.5411)	1.2738 (0.8419,1.9323)	0.1849 (0.0809,0.2976)	1.2922 (0.8253,1.9347)

	Size 4		Size 5 (Large)	
	Val-Gr	W-L	Val-Gr	W-L
Regime 1	0.6981 (0.6520,0.7559)	4.43157 (3.8702,4.799)	0.7617 (0.6450,0.8566)	4.4110 (3.7568,4.8146)
Regime 2	0.6768 (0.5581,0.8359)	3.0214 (2.2810,3.7489)	0.6217 (0.4501,0.7692)	2.8464 (2.0979,3.6757)
Diff-in-Diff	0.0174 (−0.1188,0.1412)	1.2572 (0.8220,1.8571)	0.1281 (0.0141,0.2822)	1.3442 (0.8616,2.0611)

Table 9: The first two rows report the conditional steady states for the spread in adjusted book-market ratios between the high and low return premia porfolios in each regime. The columns labeled "Val-Gr" report the spreads for portfolios sorted along the book-market dimension, in a given size category (extreme value minus extreme growth). The columns labeled "W-L" report the spreads for portfolios sorted along the recent past return performance dimension (extreme winner minus extreme loser). The row labeled "Diff-inDff" reports the difference between these spreads across the two wealth ratio/interest rate regimes. The book-market ratios are adjusted for the differences in expected earnings, so the adjusted book-market spreads coincide with spreads in the present discounted values of expected excess returns. In parentheses we report 68% posterior credible sets. The sample spans the period 1964:Q1-2013:Q3.

Probability of a decline in Premia

	Size 1 (Small)		Size 2		Size 3	
Spread	V-G	W-L	V-G	W-L	V-G	W-L
% draws	91.25	100	99.85	100	95.20	100
	Size 4		Size 5 (Large)			
Spread	V-G	W-L	V-G	W-L		
% draws	54.85	99.95	86.70	100		
Portfolio	V (S1)	V (S2)	V (S3)	V (S4)	V (S5)	Winner
% draws	99.90	99.85	99.00	94.50	95.00	77.60
Portfolio	G (S1)	G (S2)	G (S3)	G (S4)	G (S5)	Loser
% draws	94.20	6.65	49.80	88.60	88.90	0.05

Table 10: The table reports the probability of a reduction in the estimated risk premia when moving from the low market valuation regime 1 to the high market valuation regime 2. These probabilities are computed as the percentage of draws for which the premia decline. The first and second rows refer to the spread portfolios, whereas the third and fourth rows refer to the individual portfolios.

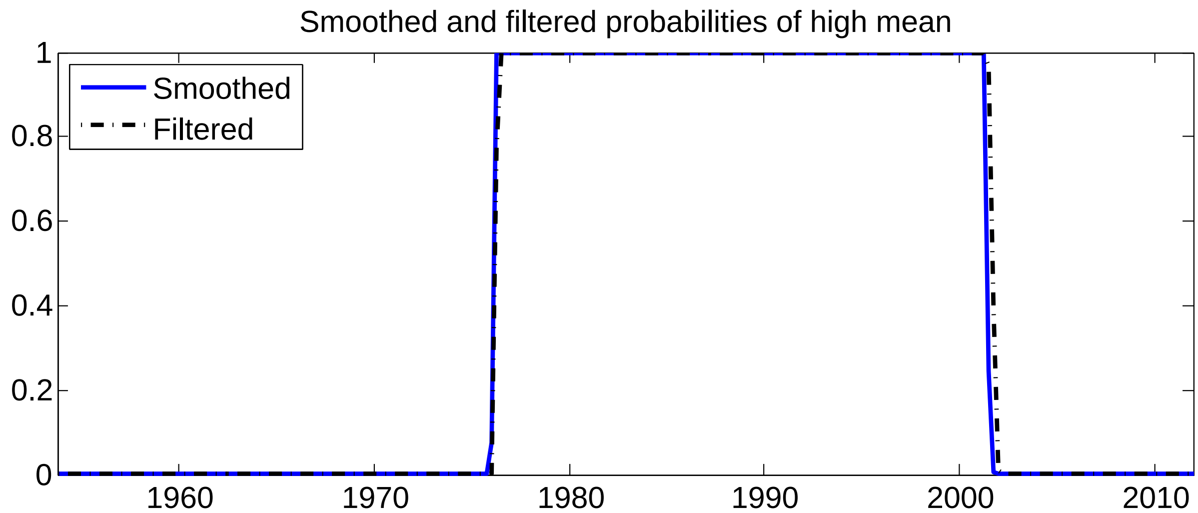

Figure 1: Smoothed probability of high mean regime for the Markov-switching cointegrating relation. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

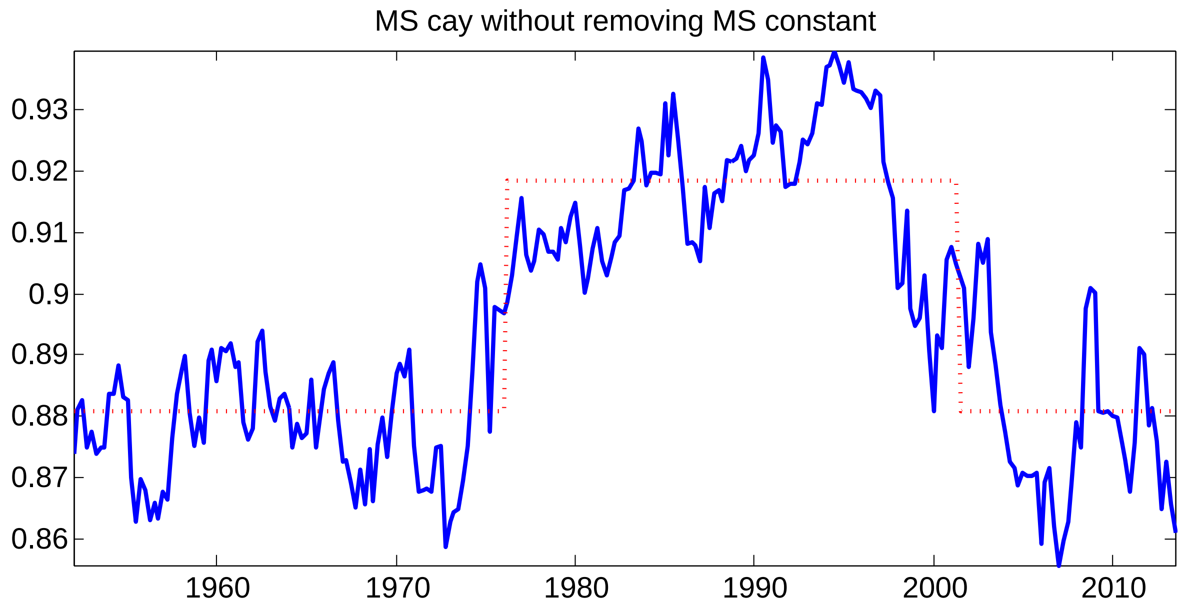

Figure 2: The Markov-switching estimated cay^{MS} is plotted with*out* removing the constant. The red dashed lines are the values of α_1 and α_2, which correspond to the most likely mean values in each regime. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

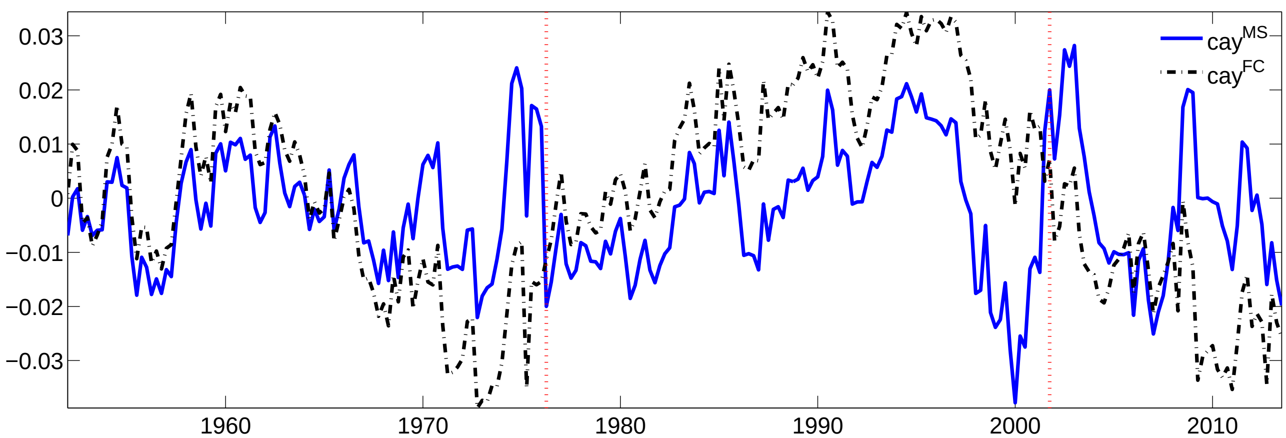

Figure 3: Markov-switching and fixed coefficients *cay*. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

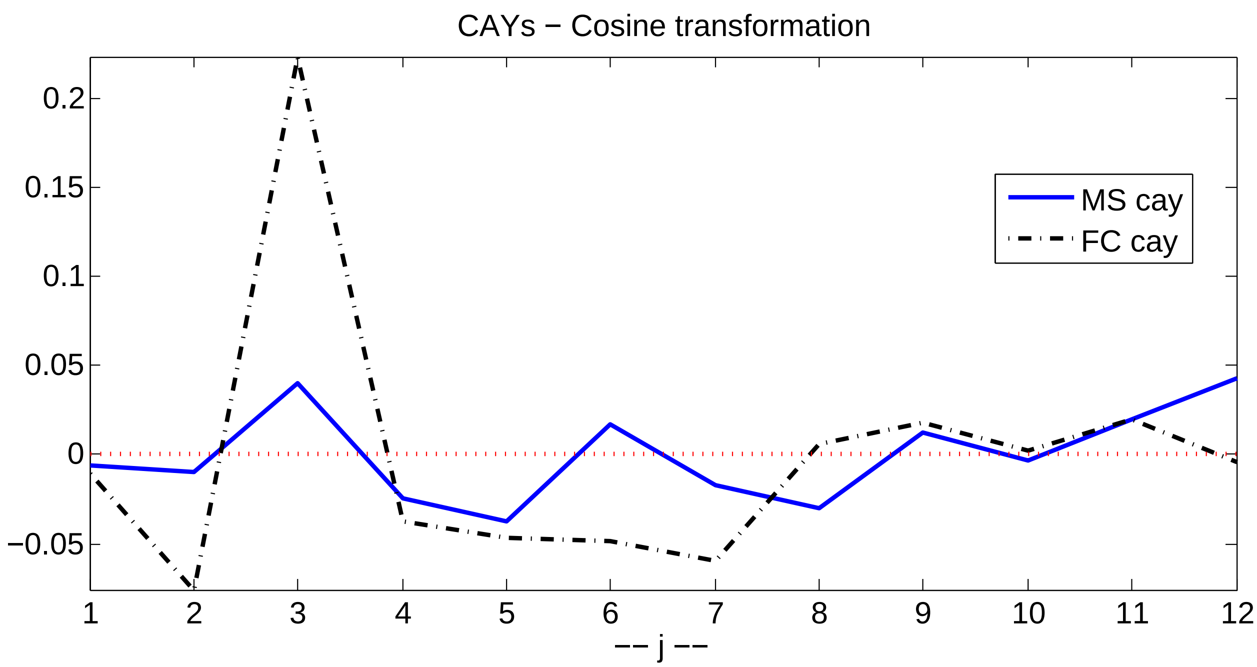

Figure 4: Low frequency averages of *cay*. The figure plots the set of averages $\{f_j\}_{j=1}^{k}$, which capture the variability in *cay* for periods greater than $2T/k$, where T is the sample size. Thus, with T $= 247$ quarters, the $k = 12$ points plotted summarize the variability in *cay* for periods greater than $2 * 247/12 = 41.1667$ quarters, approximately 10 years. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

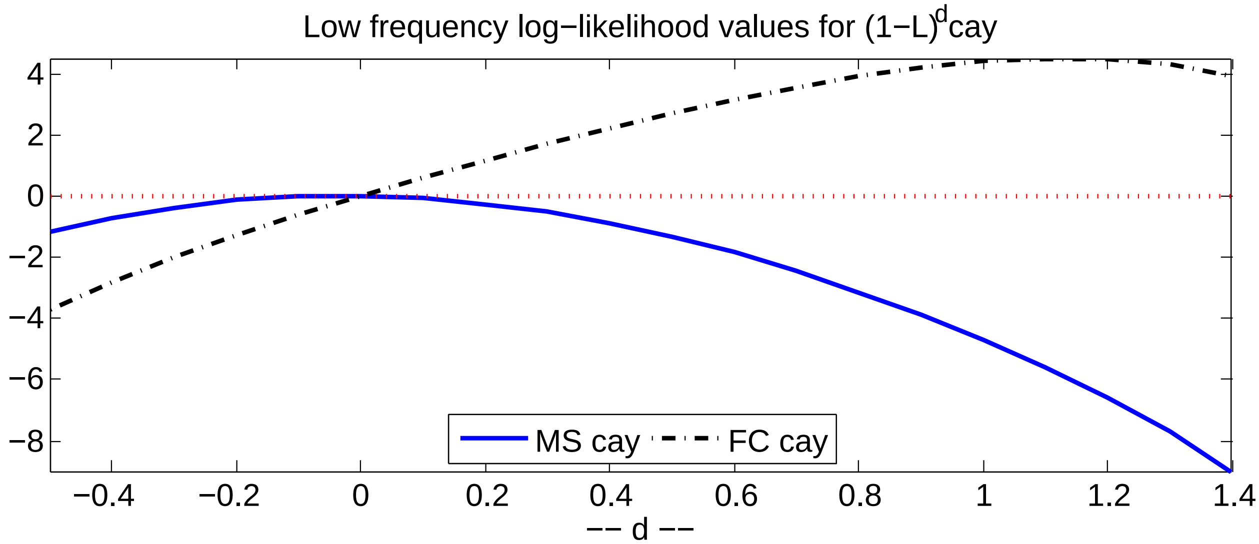

Figure 5: Low frequency log likelihood values for $(1-L)^d cay_t$. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

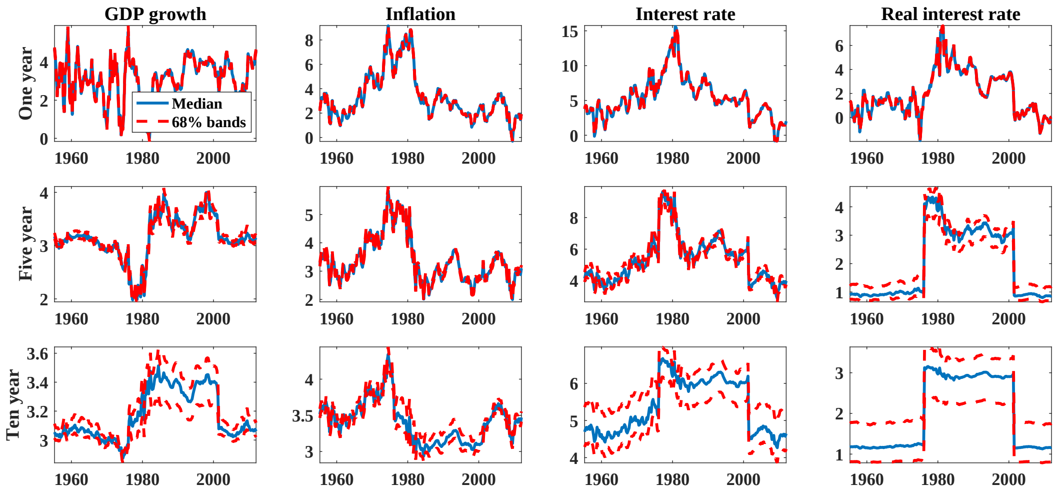

Figure 6: Projections from MS-VAR. The figure reports the conditional expectations based on the MS-VAR at different horizons taking into account the possibility of regime changes.

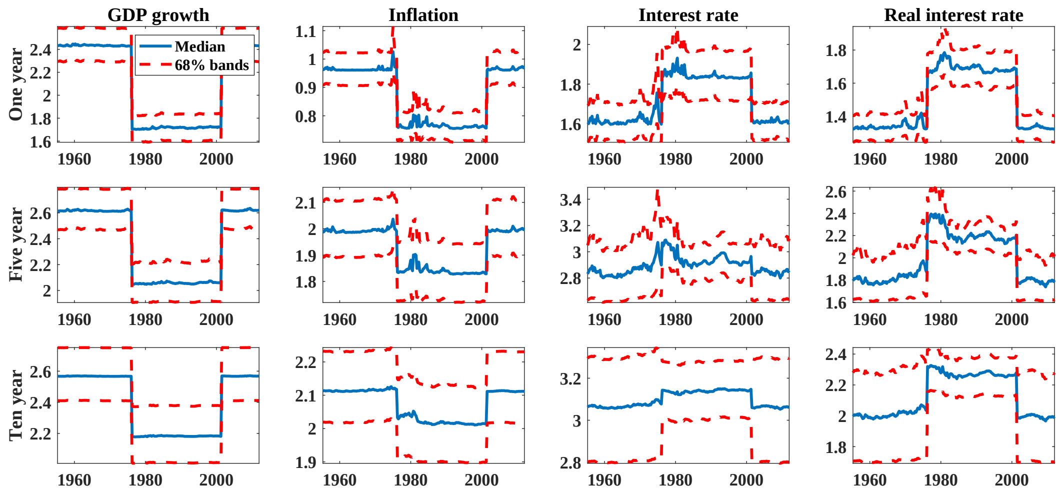

Figure 7: Uncertainty based on MS-VAR. The figure reports the conditional standard deviations at different horizons based on the MS-VAR taking into account the possibility of regime changes.

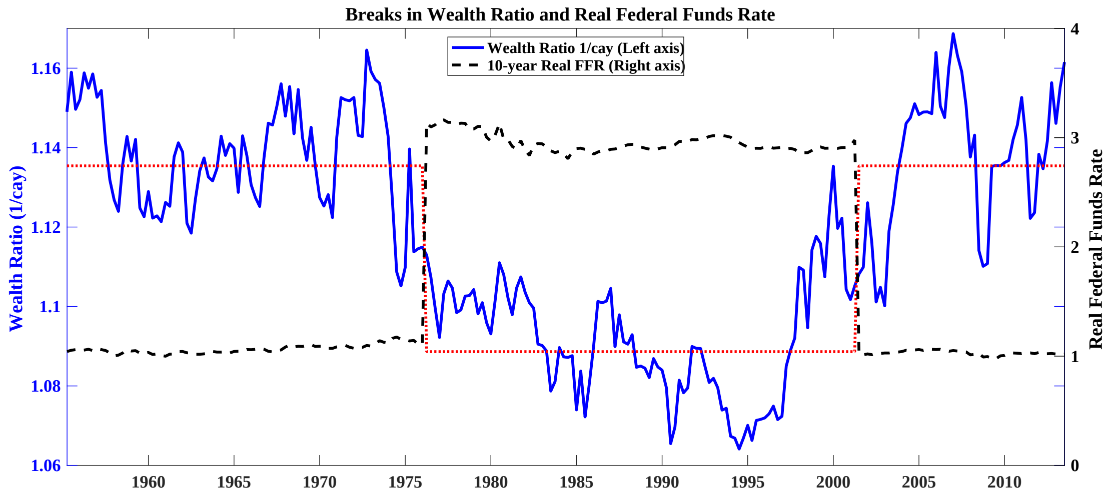

Figure 8: Wealth Ratio and federal funds rate. The wealth ratio (solid blue line, left axis) is plotted together with the ten-year-ahead real federal funds rate (black dashed line, right axis). The wealth ratio is obtained as the inverse of cay^{MS} without removing the Markov-switching constant. The red dashed line represents the inverse of the regime-probability weighted average of the constants α_1 and α_2. The ten-year-ahead real federal funds rate is computed as the ten-year-ahead expected value of the real federal funds fate as implied by the Markov-switching VAR. The sample is quarterly and spans the period 1955:Q4 to 2013:Q3. With respect to the estimates for cay^{MS} the sample is adjusted to take into account data availability for the Federal Funds rate.

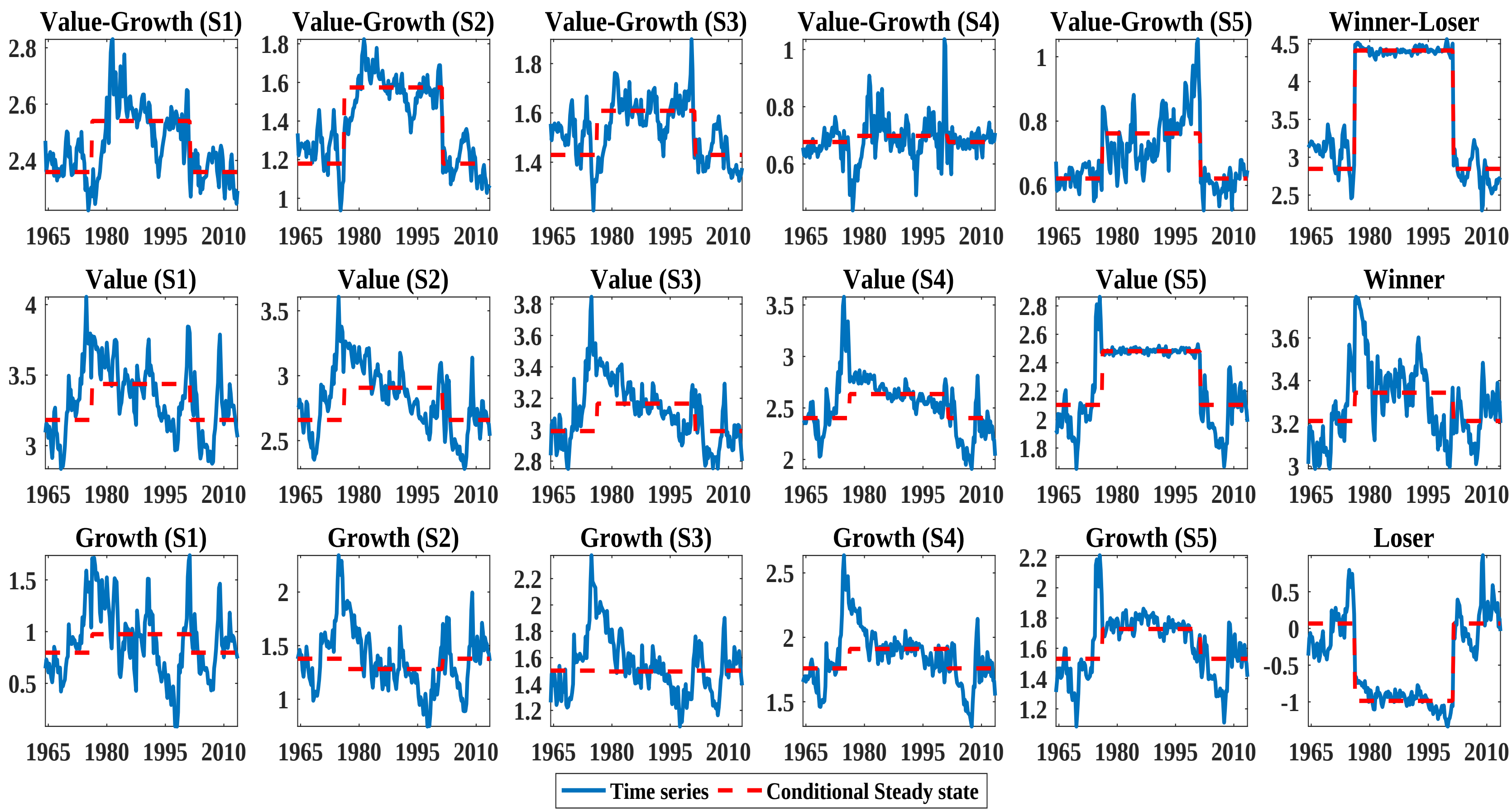

Figure 9: Evolution of Risk Premia. The figure reports the evolution of the the PDV of risk premia for different portfolios. The blue solid line reports the evolution of the risk premia over time, while the red dashed line corresponds to the conditional steady state of the PDV based on the regime in place. Both are computed by taking into account the possibility of regime changes. The sample spans the period 1964:Q1-2013:Q3.