# Monetary Policy and Asset Valuation*

Francesco Bianchi
Duke, CEPR, NBER

Martin Lettau
UC Berkeley, CEPR, NBER

Sydney C. Ludvigson
NYU and NBER

First draft: May 12, 2016
This draft: June 15, 2018

**Abstract**

We find evidence of infrequent shifts, or "regimes," in the mean of the asset valuation variable $cay_t$ that are strongly associated with low-frequency fluctuations in the real federal funds rate, with low policy rates associated with high asset valuations, and vice versa. There is no evidence that infrequent shifts to high asset valuations are associated with higher expected economic growth or lower economic uncertainty; indeed, the opposite is true. Additional evidence shows that regimes of low interest rates and high asset valuations are characterized by lower equity market risk premia and monetary policy that is less responsive to inflation.
JEL: G10, G12, G17

*Bianchi: Department of Economics, Duke, 213 Social Sciences building, Box 90097, Durham, NC 27708-0097; Email: francesco.bianchi@duke.edu, https://sites.google.com/view/francescobianchi/home. Lettau: Haas School of Business, University of California at Berkeley, 545 Student Services Bldg. #1900, Berkeley, CA 94720-1900; E-mail: lettau@haas.berkeley.edu; Tel: (510) 643-6349, https://sites.google.com/view/martinlettau/. Ludvigson: Department of Economics, New York University, 19 W.4th Street, 6th Floor, New York, NY 10012; Email: sydney.ludvigson@nyu.edu; https://www.sydneyludvigson.com. We thank Ian Dew-Becker, Emmanuel Farhi, Simon Gilchrist, Matteo Maggiori, Emi Nakamura, Monika Piazzesi, Jon Steinsson, Andrea Tamoni, Harald Uhlig, and Michael Weber for helpful comments and seminar participants at the Bundesbank, Columbia, HEC Montreal, NYU, UCL, UCLA, the University of Texas at Austin, Minneapolis Fed, Indiana University, Carey – Johns Hopkins University, the National Bank of Belgium, Booth – University of Chicago, the Workshop on Monetary Policy, NY Fed, April 2017; the Financial Markets and Macroeconomic Performance conference, Frankfurt, May 2017; the Developments in Empirical Monetary economics Conference, FRB May 2017; the SED June 2017; the CCBS Research Forum on Macrofinance, Bank of England June 2017, the Barcelona GSE Summer Symposium, June 2017; the Computing in Economics and Finance conference, June 2017, Wolfe Research's Global Quantitative and Macro Investment Conference, November 2017, and the 2018 American Economic Association meetings. We thank Paulo Manoel for excellent research assistance. Any errors or omissions are the responsibility of the authors.

# 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. One might reasonably attribute the origins of this variation to market "noise" around a more stable economic state. But what if a significant fraction of this variation is attributable to lower frequency, decades-long shifts in these relative relationships? Such a phenomenon, were it to exist, could not be readily attributed to transitory stock market noise, but would instead raise questions about the role of structural changes in the macroeconomy that govern how far and how persistently asset values can deviate from their historical relationship to measures of fundamental value.

This paper presents empirical evidence of just such a lower frequency phenomenon. Our analysis builds a novel set of statistical facts on how asset values, macroeconomic aggregates, short-term interest rates, and risk premia have comoved over longer periods of time in post-war U.S. data. Broadly speaking, these findings may be delineated into four distinct categories:

1. *Asset valuations*: The U.S. economy is characterized by large, longer-term regime shifts in asset values relative to measures of broad-based macroeconomic prosperity.

2. *Macroeconomic fundamentals*: Regimes in which asset values are persistently high relative to macroeconomic fundamentals cannot be understood as the outcome of good news about the macroeconomy. On the contrary, prolonged periods of high asset valuations coincide with persistently weak economic growth and persistently high macroeconomic volatility. The converse is true for low valuation regimes.

3. *Monetary policy*: Regimes in which asset valuations are persistently high are associated with persistently low values for the real federal funds rate, and vice versa for low valuation regimes. High asset valuation regimes are also characterized by evidence of less monetary policy activism toward inflation and greater activism toward output growth.

4. *Risk premia*: Low policy rate/high asset valuation regimes coincide with lower equity market risk premia.

To establish these statistical facts, we begin by documenting evidence of infrequent but stationary shifts, or "regimes," in the mean of the consumption-wealth variable $cay_t$ of Lettau and Ludvigson (2001) (LL hereafter), one of the broadest asset valuation metrics available relative to macroeconomic fundamentals. This variable uses data on total household net worth (asset wealth, $a$) and two key macroeconomic fundamentals: consumer spending, $c$, and labor income, $y$. In this paper we estimate a new consumption-wealth variable, denoted $cay_t^{MS}$, that

can be represented as the sum of two components: $cay_t^{MS} \equiv \alpha_{\xi_t^\alpha} + \epsilon_t^{MS}$. The first component, $\alpha_{\xi_t^\alpha}$, is a time-varying conditional mean, modeled as discrete-valued random variable that follows a regime switching process. The traditional $cay_t$ ignored time-variation in this component. Our estimates for this study imply that it varies considerably but that its fluctuations are dominated by lower frequency variation. The second component, $\epsilon_t^{MS}$, is a continuous-valued random variable whose fluctuations are comprised of higher frequency fluctuations around $\alpha_{\xi_t^\alpha}$.

The sample is divided into three clear subperiods characterized by two estimated regimes for $\alpha_{\xi_t^\alpha}$: a low asset valuation regime that prevails in a middle subperiod of the post-war era, from 1976:Q2 to 2001:Q2, and a high asset valuation regime that prevails at the beginning and end of our sample, namely 1952:Q1-1976:Q1 and the post-millennial period 2001:Q3-2013:Q3. Though persistent, the fluctuations captured by our estimated regime switches are not permanent.

With these estimated regimes in hand, we direct our attention to the main question of the study, namely what these infrequent asset valuation regimes may be linked to in the macroeconomy. To do so we estimate a Markov-switching vector autoregression (MS-VAR) in macroeconomic data. Estimates from the MS-VAR are used to compute conditional forward looking moments of the macro variables that take into account the possibility of regime changes. In doing so, we find evidence of large breaks in the conditional expected value of the 5- and 10-year-ahead real federal funds rate that coincide with the breaks in the mean of $cay^{MS}$. Regimes of low asset valuations (high $cay^{MS}$) are associated with an expectation of persistently high values for the real federal funds rate, while regimes of high asset valuations (low $cay^{MS}$) are associated with an expectation of persistently low values. The post-millennial period in particular, one characterized by high asset valuations according to a number of indicators, is marked by forecasts of prolonged low values for the real federal funds rate, in contrast to the middle subperiod where asset valuations are low and conditional expected policy rates are high. We further find that one measure of the stance of monetary policy from an estimated interest rate rule–namely the long-run response of the nominal federal funds rate to a permanent change in inflation or output growth–differs markedly across the low and high asset valuation regimes, as explained below.

One might reasonably attribute the high asset valuation regimes to more favorable prospects for economic growth and/or lower uncertainty about that growth. But this is not what we find. Indeed, the high asset valuation subperiods are times of relative weakness in GDP growth, investment growth, and R&D growth, a declining labor share, and relatively high macroeconomic volatility. These findings 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. The one exception to these findings of broad-based economic weakness during high asset valuation regimes is the stock market itself: high asset valuation subperiods are characterized by forecasts of higher dividend growth on publicly

traded shares.

These findings raise a question: why are persistent shifts in the Central Bank's conditional expected policy rate associated with such pronounced low frequency shifts in asset valuations? One possible answer that is consistent with our evidence is simply that macroeconomic fundamentals play no role in asset valuations and the results solely reflect regime shifts in discount rates. But even if we restrict attention to explanations based on changing discount rates, theories differ on *why* discount rates change with policy rates. Some theories tie low and declining discount rates entirely to endogenous movements in the so-called "natural" rate of interest with no role for monetary policy, as in the recent strand of models for which low interest rates are the result of shocks that increase the fraction of wealth held by 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. By contrast, other theories imply that shifts downward in the risk-free rate coincide with shifts downward in risk premia, as in those models that can be broadly characterized as having a "reaching for yield" channel (e.g., Rajan (2006); Rajan (2013); Drechsler, Savov, and Schnabl (2014); Piazzesi and Schneider (2015); Acharya and Naqvi (2016); Coimbra and Rey (2017)). These theories often ascribe a role for monetary policy.

The contradictory implications for risk premia under these two frameworks provides a testable implication. We present evidence from equity markets that the low interest rate regimes we document coincide with lower equity market risk premia, consistent with reaching for yield. Specifically, we find that, in a switch from a high to low interest rate regime, the estimated present discounted value of future risk premia on the aggregate stock market, as well as that of several other equity portfolios, simultaneously fall to lower levels. Moreover, the risk premia and the book-market ratios (adjusted for expected earnings) of evidently riskier, higher Sharpe ratio portfolios, such as those that go long in value stocks or stocks that have recently appreciated the most, fall 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. The estimated risk premia on all of these equity portfolios 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)). The evidence in these papers pertains to heavily intermediated asset classes. By contrast, our evidence pertains to equity market portfolios, an asset class ostensibly held by retail investors and households, as well as intermediaries.

We present evidence of a role for systematic monetary policy in these outcomes. Specifically, we find that the Taylor rule implied by our MS-VAR differs significantly across the two regimes, with the high interest rate/low asset valuation regime characterized by a funds rate that is more responsive to inflation, and the low interest rate/high asset valuation regime characterized by a funds rate that is more responsive to output growth and much less responsive to inflation. We argue below that our findings on asset valuations, risk premia, and policy rates are not readily explained as the outcome of fluctuations in the natural rate alone, or (equivalently) as the result of a policy rate that perfectly tracks the natural rate over time.

A survey of the asset pricing literature over the past several decades reveals years of empirical evidence that risky financial valuation ratios such as $cay_t$ are dominated by fluctuations in risk premia and covary much less strongly, if at all, with interest rates. From this perspective, the results of this study are quite surprising. The avenue by which such variation is uncovered in this paper is the specification of a Markov-switching version of the variable with two distinct components, one that covaries strongly with both short-term interest rates and equity market risk premia at low frequencies, and another, $\epsilon_t^{MS}$, that predicts excess stock market returns at higher frequencies. The penultimate section of the paper presents a decomposition of variance that demonstrates the importance of both components to the overall variation in $cay_t^{MS}$, and shows that explicitly distinguishing the two renders $cay_t^{MS}$ a superior predictor of excess stock market returns compared to the traditional $cay_t$ where these components are conflated.

A growing empirical literature documents a linkage between monetary policy surprises 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)) or by documenting that a VIX-implied measure of risk aversion declines in response to expansionary monetary policy shocks (Bekaert, Hoerova, and Duca (2013)). These higher frequency analyses contrast with the evidence of this paper, which implies that lower frequency structural shifts in equity market return premia and asset values are associated with persistent changes in both the Federal Reserve's primary policy instrument and in the parameters of an estimated systematic monetary policy rule.

Our empirical findings are broadly consistent with a theoretical literature of intermediary risk-taking behavior prompted by very accommodative or unconventional monetary policies (see for example, Diamond and Rajan (2012); Farhi and Tirole (2012); Rajan (2013); Stein (2013)). These theories are closely related to those referenced above that have a reaching for yield component.

The rest of the paper is organized as follows. The next section discusses the estimation of a

Markov-switching $cay_t$ and results from this estimation. Section 3 discusses estimation of the Markov-switching VAR in macroeconomic data, focusing on evidence for what may be behind the regimes in asset valuations. This section also discusses how our findings relate to systematic monetary policy and the natural rate of interest. Section 4 investigates whether the high asset valuation/low interest rate regimes are characterized by lower risk premia in equity market assets. Section 5 considers the role of $\epsilon^{MS}$ in stock market risk premia variation and presents a variance decomposition of $cay^{MS}$ into fractions attributable to its two components. Section 6 briefly remarks on the behavior of other stock market valuation ratios over our estimated regime subperiods and Section 7 concludes. A large amount of additional material, test results, and a detailed data description have been placed in an Appendix for online publication.

# 2 Econometric Model of $cay^{MS}$

This section describes an econometric model of regime switches in the mean of $cay_t$. Under assumptions described in LL and elaborated on in Lettau and Ludvigson (2010), $cay_t$ is derived from an approximate formula for the log consumption- aggregate (human and non-human asset) 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 $y_t$ can be motivated from an aggregated household budget constraint. Specifically, if labor income is the dividend paid to human capital, one can derive an approximate expression linking $c_t$, $a_t$, and $y_t$ to expected future returns to asset wealth, consumption growth, and labor income growth:

$$cay_t \equiv c_t - \gamma_a a_t - \gamma_y y_t \approx \alpha + \mathbb{E}_t \sum_{i=1}^{\infty} \rho_w^i \left( (1-\nu) r_{a,t+i} - \Delta c_{t+i} + \nu \Delta y_{t+1+i} \right),$$

where $\nu$ is the steady state ratio of human wealth to asset wealth and $r_{a,t}$ is the log return to asset (non-human) wealth. Theory typically implies that $c_t$, $a_t$, and $y_t$ should be cointegrated, or that $cay_t$ should be covariance stationary.

In the standard estimation without regime shifts in any parameters, the stationary linear combination of $c_t$, $a_t$, and $y_t$ may be written

$$cay_t^{FC} \equiv c_t - \gamma_a a_t - \gamma_y y_t = \alpha + \epsilon_t^{FC}, \tag{1}$$

where the parameters to be estimated are $\alpha$, $\gamma_a$, and $\gamma_y$. The residual $\epsilon_t^{FC}$ is the mean zero stationary linear combination of these data, referred to as the cointegrating residual. The superscript "$FC$" stands for "fixed coefficients" to underscore the fact that no parameters are time-varying.

In this paper, we estimate a Markov-switching version of this variable, analogously written as

$$cay_t^{MS} \equiv c_t - \beta_a a_t - \beta_y y_t = \alpha_{\xi_t^\alpha} + \epsilon_t^{MS}. \tag{2}$$

The intercept term, $\alpha_{\xi_t^\alpha}$, is a time-varying mean that depends on the existence of a latent state variable, $\xi_t^\alpha$, presumed to follow a two-state Markov-switching process with transition matrix $\boldsymbol{H}^\alpha$. Thus $\alpha_{\xi_t^\alpha}$assumes one of two discrete values, $\alpha_1$ or $\alpha_2$. The choice of two regimes is not crucial, but provides an easily interpretable way to organize the data, into a low and a high valuation regime. The residual $\epsilon_t^{MS}$ is a stationary continuous-valued variable by assumption. The slope coefficients $\beta_a$ and $\beta_y$ are analogous to $\gamma_a$ and $\gamma_y$ in the fixed coefficient regression (1). They are denoted differently to underscore the point that the coefficients in (1) and (2) are not the same, just as the parameters $\alpha$ and $\alpha_{\xi_t^\alpha}$, and the residuals $\epsilon_t^{FC}$ and $\epsilon_t^{MS}$ are not the same. Because our procedure jointly recovers the slope coefficients $\beta_a$ and $\beta_y$, the timing of regime changes, and, as an implication, the decomposition of $cay_t^{MS}$ into $\alpha_{\xi_t^\alpha}$ and $\epsilon_t^{MS}$, all three statistical objects can differ. We denote the set of parameters to be estimated collectively as $\boldsymbol{\theta}$.

We use Bayesian methods with flat priors to estimate the model parameters in (2) over the period 1952:Q1-2013:Q3. The sequence $\xi^{\alpha,T} = \{\xi_1^\alpha, ..., \xi_T^\alpha\}$ of regimes in place at each point is generally unobservable and needs to be inferred jointly with the other parameters of the model. An estimate of $\alpha_{\xi_t^\alpha}$ is formed by weighting its two estimated values by their state probabilities at each point in time. For this purpose, we consider two estimates of the state probabilities. Let $T$ be the sample size used in the estimation. Let $P\left(\xi_t^\alpha = i|\boldsymbol{Z}_t;\boldsymbol{\theta}\right) \equiv \pi_{t|t}^i$ denote the probability that $\xi_t^\alpha = i$, for $i = 1, 2$, 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 that can be extracted from the whole sample: $P\left(\xi_t^\alpha = i|\boldsymbol{Z}_T;\boldsymbol{\theta}\right) \equiv \pi_{t|T}^i$. Thus, $cay_t^{MS}$ may be decomposed one of two ways:

$$cay_t^{MS} = c_t - \left(\beta_a a_t + \beta_y y_t\right) = \overline{\alpha}_t^{filt} + \epsilon_t^{MSfilt}. \tag{3}$$

$$cay_t^{MS} = c_t - \left(\beta_a a_t + \beta_y y_t\right) = \overline{\alpha}_t + \epsilon_t^{MS}. \tag{4}$$

where $\overline{\alpha}_t^{filt} = \sum_{i=1}^2 \pi_{t|t}^i \alpha_i$ and $\overline{\alpha}_t = \sum_{i=1}^2 \pi_{t|T}^i \alpha_i$ are the weighted averages of the Markov-switching means using filtered and smoothed probabilities, respectively. Note that while $cay_t^{MS}$ is the same no matter which method is used to compute regime probabilities, its decomposition differs depending on the way the two regime values are weighted.

The posterior distribution of the empirical model (2) and the corresponding regime probabilities $\pi_{t|t}^i$ and $\pi_{t|T}^i$ are obtained by computing the likelihood 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 $cay_t^{MS}$ and its decomposition into $\overline{\alpha}_t$ and $\epsilon_t^{MS}$ are based on the posterior mode of the parameter vector $\boldsymbol{\theta}$ and the corresponding regime probabilities. Uncertainty about the parameters, or about any transformation of the model parameters, is characterized using a Gibbs sampling algorithm. The full statement of the procedure and sampling algorithm is given in the Appendix.

The variables $cay_t^{FC}$ and $cay_t^{MS}$ may be interpreted as log inverse asset valuation ratios, akin

to a log dividend-price ratio as opposed to log price-dividend ratio. For simplicity, we sometimes refer these as inverse "wealth" ratios, even though they are not the same as measures of the inverse consumption-aggregate wealth ratio.[1] Since, in population, $\epsilon_t^{MS}$ and $\epsilon_t^{FC}$ are mean zero random variables, the intercept terms give the mean of these inverse asset valuation ratios.[2] A high $\alpha_i$ corresponds to a low mean 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 $y_t$. We refer to $cay_t^{MS}$ interchangeably as the inverse wealth ratio, or equivalently define the (log) *wealth ratio* as $-cay_t^{MS} = -\left[\epsilon_t^{MS} + \overline{\alpha}_t\right]$, where we use the smoothed regime probability decomposition as our benchmark.

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. Model comparison tests based on the $BIC$ criterion show that $cay_t^{MS}$ describes the data far better than a model with fixed coefficients and no switches in the constant. These results are presented in the "Model Comparison" section, Appendix 7, of the Appendix.

The results show that the sample is divided into three clear subperiods characterized by the two regimes for $\alpha$. Regime 1 is a high $\alpha$ regime with the posterior mode point estimate equal to $\hat{\alpha}_1 = 0.9186$. The low $\alpha$ regime 2 posterior mode estimate is $\hat{\alpha}_2 = 0.8808$. A high $\alpha$ regime for *cay* corresponds to a low valuation ratio for the stock market, analogous to a low price-dividend ratio. We therefore refer to *high $\alpha$ regime 1* as the *low asset valuation regime*, and *low $\alpha$ regime 2* as the *high asset valuation regime*. Figure 1 shows that the low asset valuation regime prevails for a prolonged period of time from 1976:Q2 to 2001:Q2, during which the smoothed probability that $\alpha = \hat{\alpha}_1$ is remains close to unity. By contrast, the pre-1976 and post-2001 subsamples are high asset valuation regimes, where the probability that $\alpha = \alpha_1$ is virtually 0. These correspond to the subperiods 1952:Q1-1976:Q1, and 2001:Q3-2013:Q3, respectively. The filtered and smoothed probabilities are found to be very similar and there is little uncertainty about the timing of regime changes. (The Appendix explains why this is the case.) For the rest of the paper we focus primarily on $cay_t^{MS}$, the version of Markov-switching *cay* obtained using smoothed probabilities.

Table 1 reports summary statistics for the difference $\hat{\alpha}_1 - \hat{\alpha}_2$, along with percentiles of the posterior distribution for $\hat{\alpha}_1 - \hat{\alpha}_2$. The 90% credible set for $\hat{\alpha}_1 - \hat{\alpha}_2$ is non-zero and positive,

[1] Although $cay_t$ is derived from an approximate formula for the log consumption-aggregate (human plus nonhuman) wealth ratio $c_t - w_t$, it is proportional to it only under certain special circumstances. (See Lettau and Ludvigson (2010) for a discussion.) Lustig, Van Nieuwerburgh, and Verdelhan (2013) point out that it is possible to directly measure $c_t - w_t$ if one is willing to take a stand on a specification for the stochastic discount factor. Doing so, they provide an estimate of $c_t - w_t$ that behaves differently from the *cay* measures of this paper, primarily because it does not ascribe as important of a role to time-varying equity market risk premia in its variation. We discuss this further below.

[2] In a finite sample, $\epsilon_t^{MS}$ and $\epsilon_t^{FC}$ are not necessarily mean zero because of the leads and lags of the first differences included in the DLS regression used to correct for finite sample biases–see the Appendix. In population these variables are mean-zero by definition.

indicating that the difference between the means of the inverse wealth ratio across the regimes is positive and with all posterior probability mass away from zero. The two regimes are stationary but persistent, as by the estimated diagonal elements of the transition matrix $H^{\alpha}$, also reported in Table 1. To give a visual impression of the properties of these regimes, Figure 2 plots $cay_t^{MS}$ over time. 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 inverse 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.

# 3 Macroeconomic Origins of Asset Valuation Regimes

For the rest of this paper, we search for empirical explanations from the macroeconomy for the estimated regimes observed in the mean of $cay^{MS}$. To do so we estimate a Markov-switching vector autoregression in macroeconomic data. To allow for structural change, the parameters of the MS-VAR are permitted to potentially undergo structural shifts. Importantly, however, we impose the estimated regime sequence for the conditional mean of $cay^{MS}$ on the MS-VAR while allowing the parameters characterizing the different regimes as well as the transition matrix to be freely estimated.[3] 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 objective here is not to estimate independent regimes for the variables in the MS-VAR and see if they align with the previously estimated breaks in *cay*. Instead, the objective is to establish what, if anything, is different in the MS-VAR variables across the two *previously estimated* asset valuation regimes that could help explain the breaks in the mean of *cay*. We therefore deliberately "tie our hands" by forcing the regime sequence for the MS-VAR to correspond to breaks in $\alpha_{\xi_t^{\alpha}}$. Note that there is no implication in this procedure that the macro variables must necessarily show evidence of structural change. All parameters other than the regime sequence as well as the conditional moments of the macro variables in the MS-VAR are freely estimated and could in principle show no shift across the asset valuation regimes.

All MS-VARs estimated in this section and the next are implemented using Bayesian methods with flat priors, while uncertainty about the parameters is characterized by Gibbs sampling. The Appendix provides estimation details.[4]

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

$$Z_t = c_{\xi_t} + A_{1,\xi_t} Z_{t-1} + A_{2,\xi_t} Z_{t-2} + V_{\xi_t} \varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{5}$$

[3] We choose the 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.

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

where $Z_t$ is an $n \times 1$ vector of variables, $c_{\xi_t}$ is an $n \times 1$ vector of constants, $A_{l,\xi_t}$ 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 shocks $\varepsilon_t$. The process $\xi_t$ 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 $cay^{MS}$.

In our baseline macro MS-VAR, the vector $Z_t$ includes five variables at quarterly frequency: GDP growth, inflation, investment growth, R&D growth, and the effective federal funds rate (FFR). Inflation is defined as the year-to-year differences of the logarithm of the GDP price deflator. GDP growth, investment growth, and R&D growth are defined as the year-to-year differences of the logarithm of real GDP per capita, real investment per capita, real R&D per capita, respectively. The quarterly FFR is obtained by taking the average of monthly figures of the effective federal funds rate. In a secondary macro MS-VAR, the vector $Z_t$ includes five variables at quarterly frequency: GDP growth, inflation, the change in the labor compensation share of GDP ("labor share" for short), aggregate dividend growth for all firms traded on NYSE, NASDAQ and AMEX, and the effective federal funds rate. The Data Appendix provides a detailed description of our data and sources. The sample for this estimation spans the period 1955:Q3-2013:Q3.[5]

We use the MS-VAR parameter estimates to compute the conditional expectation and the conditional standard deviation of each variable in the MS-VAR and the real interest rate (RIR), defined as the difference between the FFR and one-step-ahead conditional expectation of inflation given by the MS-VAR. 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. For each variable $z_t \in Z_t$, the conditional expectation and conditional standard deviation are given by $\mathbb{E}_t(z_{t+s})$ and $sd_t(z_{t+s}) = \sqrt{\mathbb{V}_t(z_{t+s})} = \sqrt{\mathbb{E}_t\left[z_{t+s} - \mathbb{E}_t(z_{t+s})\right]^2}$, where $\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot|\mathbb{I}_t)$ 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. Given the estimated posterior distribution of MS-VAR parameters and transition matrix $\boldsymbol{H}^A$, we compute posterior distributions for the conditional moments $\mathbb{E}_t(z_{t+s})$ and $sd_t(z_{t+s})$. The statistic $sd_t(z_{t+s})$ is a statistical measure of economic uncertainty, as implied by the MS-VAR. We discuss the results of this estimation in subsections, beginning with a focus on monetary policy and moving on to the results for macroeconomic fundamentals across our estimated regimes. The full estimation procedure is described in the Appendix.

[5]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.

## 3.1 Asset Valuations and the Federal Funds Rate

Figure 3 reports the 5- and 10-year-ahead conditional expected values of each variable in the baseline MS-VAR plus the RIR. The figure reports the median and 68% credible sets from the posterior distribution of these conditional expectations.

We begin by inspecting the behavior of the real FFR (RIR) displayed in the far right-hand column of Figure 3. There are distinct interest rate regimes over the post-war sample that line up with the asset valuation regimes. In particular, the right-most column shows evidence of significant structural change in the 5– and 10-year-ahead conditional expected values of the RIR that coincide with the regime sequence estimated for $\alpha_{\xi_t^\alpha}$. The occurrence of the low asset valuation regime in the middle subsample from 1976:Q2-2001:Q2, coincides with a sharply higher expected value 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 a much lower expected value for the RIR. The differences across subsamples are strongly statistically significant according to the 68% posterior credible sets for the conditional expected values. Because the MS-VAR parameters are freely estimated, the procedure was free to find 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. That the estimated regime shifts show up prominently in the conditional expected value of the funds rate five to ten years ahead illustrates the extent to which shifts in $\alpha_{\xi_t^\alpha}$ coincide with a persistent low or high interest rate environment, rather than transitory movements in these rates.

There is no clear pattern with inflation across the regimes. Thus the breaks in the expected real interest rate five or ten years ahead coincide with breaks in the conditional expected value of the nominal interest rate, which the Federal Reserve directly influences. The Federal Reserve may also directly influence expected inflation. But movements in the conditional expectation of inflation do not line up well with the regime sequence for breaks in $\alpha_{\xi_t^\alpha}$. 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.

Figure 4 gives a visual impression of the inverse relation between the persistent changes in the RIR and asset valuations. The figure plots the log "wealth ratio" (the negative of $cay^{MS}$), along with the ten-year-ahead conditional expected value of the real federal funds rate implied by the baseline MS-VAR, 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 $-\overline{\alpha}_t$). There are clear regime shifts in wealth ratio that move from high-low-high over the post-war sample,

coinciding with a low-high-low expected value for the 10-year-ahead real federal funds rate. The magnitude of the regime shifts in the expected funds rate are large, ranging from about 1% in the high asset valuation/low interest rate regimes, to 3% in the low asset valuation/high interest rate regime.

Figure 5 superimposes the five- and ten-year-ahead expected real FFR implied by the MS-VAR with the *quarterly* real FFR, defined as the quarterly nominal FFR minus the one-step-ahead expected inflation rate implied by the MS-VAR. For comparison, the dashed line computes the quarterly real FFR using survey expectations of inflation, equal to the nominal FFR minus the median forecast of CPI inflation one quarter ahead from the Survey of Professional Forecasters (this series starts in 1983).[6] The quarterly real FFR is of interest because it is policy tool under the direct control of the Central Bank. Figure 5 that there have been persistent shifts the quarterly real funds rate that caused it to fluctuate around very different levels across our regimes, with low values in the early and late subperiods and high values in the middle subperiod. The low-high-low pattern in the MS-VAR estimates of the five- and ten-year ahead conditional expected values for the real funds rate are merely a reflection of these persistent mean shifts in the quarterly funds rate. Although the shocks associated with the persistently low interest rate policies in the early and late subperiods were undoubtedly different, the two subperiods are nonetheless characterized by similar outcomes for asset valuations and (as we show later) risk premia.

## 3.2 Systematic Monetary Policy

We consider the possible role of secular changes in the stance of systematic monetary policy in the lower frequency asset valuation and interest rate regimes documented above. We do not consider the role of a monetary policy "shock," i.e. an innovation in a Taylor-type policy rule (Taylor (1993)), because such shocks are less likely to have long-lasting effects on real interest rates than are changes in the stance of monetary policy.[7] Instead, we estimate the Central Bank's systematic reactions to economic fluctuations, part of its deliberate policymaking, that are more likely to be relevant for our lower frequency findings. Shifts in systematic monetary policy could occur, for example as part of the Central Bank's judgement over how long and how extensively to accommodate a persistent economic disturbance.

To estimate the stance of monetary policy, we compute the VAR-implied Taylor rule allowing

[6] The survey forecast is extremely volatile.We therefore use a four quarter moving average of the median forecast of one-quarter-ahead inflation.

[7] Christiano, Eichenbaum, and Evans (2005) estimate that Taylor-rule innovations have short-run effects on real interest rates (but see Miranda-Agrippino and Rey (2015) for more recent evidence suggesting financial returns are affected more persistently by such innovations.) Changes in the weights on inflation or output growth or changes in the inflation or output target, have been found to have both larger and more long-lasting effects on the real interest rate than an innovation in a Taylor-type rule in DSGE models (Rudebusch and Swanson (2012); Kliem, Meyer-Gohde, and Moyen (2017)).

the parameters of this rule to potentially vary across the two previously estimated regimes. Following Primiceri (2005) and Sims and Zha (2006), we use the approach to characterize the stance of monetary policy in each regime, specifically the degree of activism in the systematic monetary policy responses to macroeconomic objectives. The goal is to compare these stances across regimes.

Consider the MS-VAR equation describing the behavior of the FFR. Equation (5) describes the reduced-form VAR, but here we focus on an implied structural-form. Zeroing in on the terms relevant for the responses of the FFR to inflation and output growth (i.e., ignoring all other terms), the structural-form equation for the FFR is:

$$\begin{aligned} FFR_t &= \psi_{0,\pi,\xi_t}\pi_t + \psi_{1,\pi,\xi_t}\pi_{t-1} + \psi_{2,\pi,\xi_t}\pi_{t-2} \\ &\quad + \psi_{0,\Delta GDP,\xi_t}\Delta GDP_t + \psi_{1,\Delta GDP,\xi_t}\Delta GDP_{t-1} + \psi_{2,\Delta GDP,\xi_t}\Delta GDP_{t-2} \\ &\quad + \psi_{1,FFR,\xi_t}FFR_{t-1} + \psi_{2,FFR,\xi_t}FFR_{t-2} + \omega_{FFR,t} \end{aligned} \tag{6}$$

where the parameters $\psi_{\cdot,\pi,\xi_t}$ and $\psi_{\cdot,\Delta GDP,\xi_t}$ capture the contemporaneous and lagged response of the FFR to inflation and output growth, $\psi_{1,FFR,\xi_t}$ and $\psi_{2,FFR,\xi_t}$ control the persistence in the response of the FFR, and $\omega_{FFR,t}$ is a structural monetary policy shock. The expression (6) can be interpreted as a Taylor rule in which the FFR today depends on inflation, real activity, and past values of the FFR. The coefficients in (6) are obtained by multiplying the parameters obtained for the reduced-form MS-VAR (5) by an estimated rotation matrix that orthogonalizes the innovations.[8] The Appendix provides a detailed mapping between the VAR coefficients and the structural Taylor rule presented in (6).

This framework may be used to study the long run responses of the FFR to permanent increases in inflation or output growth. Specifically, suppose that the inflation rate increases permanently by 1%. Then the long term response of the FFR under regime $\xi_t$ is given by:

$$LR_{\pi,\xi_t} = \left[1 - \sum_{j=1}^{2} \psi_{j,FFR,\xi_t}\right]^{-1} \sum_{j=0}^{2} \psi_{j,\pi,\xi_t}.$$

Similarly, if we are interested in the long run response of the FFR to a permanent 1% increase in output growth, we have:

$$LR_{\Delta GDP,\xi_t} = \left[1 - \sum_{j=1}^{2} \psi_{j,FFR,\xi_t}\right]^{-1} \sum_{j=0}^{2} \psi_{j,\Delta GDP,\xi_t}.$$

Table 2 reports the results for the long term responses of the FFR to inflation and output growth. The table reports the median and 68% posterior credible sets from the posterior distribution of the long term responses. No matter which regime we condition on, a permanent

[8] We estimate our baseline VAR and use a Cholesky identification scheme to pin down the contemporaneous effects of inflation on the FFR, under the assumption that the FFR can react contemporaneously to all other variables in the VAR, while the other variables react with a lag to movements in the FFR. This identification assumption is quite common in the structural VAR analysis. See the Appendix for further discussion.

increase in inflation or GDP growth increases the FFR in the long run. But the long term response of the FFR to an increase in inflation is substantially larger under the low valuation/high real interest rate regime than it is under the high valuation/low real interest rate regime. The opposite is true for the long term response of the FFR to output growth across the regimes. In other words, under the low valuation regime, the Federal Reserve seems to be more concerned with inflation stabilization, while under the high valuation regime it seems more concerned with output stabilization. This reflects the finding that the weights on inflation and output differ significantly across the two regimes. According to the 68% credible sets, these differences across regimes are strongly statistically significant for inflation, but not for output growth.

Some shifts in the monetary policy stance captured by our estimates coincide with well worn narratives about monetary policy that are related to the unique personalities, perspectives, and circumstances of the leaders who chaired the Federal Reserve Board at the time. These narratives are labeled in quotation in Figure 5.

The first manifestation of the high asset valuation regime 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)). We refer to this as the "Burns Accommodation," after Arthur Burns who chaired the Federal Reserve Board over much of this subperiod.

The occurrence of the low asset valuation regime, in the middle subperiod from 1976:Q2-2001:Q2, is associated with the end of the Great Inflation. The beginning of this subperiod precedes Volcker's appointment as Chairman of the Federal Reserve in August 1979 and the disinflation that followed by three years. Several papers place the beginning of the Volcker disinflation in the early 1980s (e.g., Bianchi (2013) and Baele, Bekaert, Cho, Inghelbrecht, and Moreno (2015)). However, Sims and Zha (2006) estimate a structural MS-VAR and find evidence of a change in the conduct of monetary policy from less to more active on inflation toward the end of 1977, in line with the break date estimated here. Real interest rates increased significantly during the "Volcker disinflation" and remained high for a prolonged period of time, well after inflation came down. For example, by the third quarter of 1984 inflation had declined to 3.4% from where it previously stood in double digits, yet the nominal funds rate remained 11.4%. One interpretation of the persistently high real rates during this time is that they were engineered to build credibility for lower inflation. The opposite goal–building credibility for higher inflation and faster growth might be labeled an objective of the post-millennial period, especially after the Great Recession. The assumption that the Central Bank can (and in some cases should) attempt to influence real rates for extended periods underlies the rationale and

primary objective of forward guidance, which is to re-anchor expectations around a new future path for inflation and therefore real rates (e.g., Eggertsson and Woodford (2003)). There is some evidence that expectations of inflation may in fact be slow to anchor around a new level. Rossi and Sekhposyan (2016) find that Fed Greenbook and other professional survey forecasts of inflation were persistently biased downward in the 1970s and persistently biased upward in the 1980s and 1990s.[9] This is consistent with the idea that Volcker, and later Greenspan, may have felt compelled to keep real rates high for an extended period because expectations of long-run inflation were slow to come down even after inflation itself did.

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. The steady downward march in real policy rates that then took hold began three years earlier with the so-called "Greenspan Put," a perceived attempt of Chairman Greenspan to prop up securities markets after the collapse of Long Term Capital Management in 1998 by lowering interest rates and (allegedly) resulting in a perception of put protection on asset prices. 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 LTCM collapse, 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, resumed their upward march in 2009. The period of high equity valuations persists at the end of our sample with rates at or near the zero lower bound (ZLB), coinciding with the explicit forward guidance "low-for-long" policies under Chairman Bernanke that promised in 2011 to keep interest rates at ultra low levels for an extended period of time, possibly longer than that warranted by a 2% inflation objective.

The three distinct $cay^{MS}$ 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 $cay^{MS}$, 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. Taken together, the results are suggestive of a role for the Federal Reserve in driving persistent movements in equity and interest rate behavior.

[9]Figure 5 shows that the VAR estimate of the RIR and the estimate based on one-quarter-ahead concensus inflation forecast from the SPF are fairly similar, while Rossi's evidence of bias is for longer periods of time, e.g., ten years. These two findings can be reconciled by noting that small but persistent deviations from rational forecasts cumulate over time.

### 3.3 What About the Natural Rate?

Particularly since the financial crisis of 2007-2009, much attention has been given to the so-called natural rate of interest, a model-specific theoretical construct that gives a value for the short rate that would be consistent with a frictionless environment and full employment in the absence of monetary policy. One idea is that the immediate aftermath of the crisis was characterized by low and negative values for the natural rate, driven by adverse liquidity shocks in short-term credit markets that resulted in an increase in the convenience yield on safe assets (e.g., Kurlat (2013); Bigio (2015); Del Negro, Eggertsson, Ferrero, and Kiyotaki (2017)). Other factors that could lead to low or negative values for the natural rate in economic models are adverse shocks to productivity and/or uncertainty. These are all plausible characteristics of the financial crisis and its immediate aftermath. We argue here, however, that it is more challenging to explain our findings on asset valuations and policy rates over longer periods of time as the outcome of fluctuations in the natural rate alone, or (equivalently) as the result of a policy rate that perfectly tracks the natural rate over time. We make this argument for several reasons.

First, although macroeconomic theories on the endogenous determination of the natural rate don't typically derive implications for asset prices or risk premia, it is doubtful that the forces that drive the natural rate to be low or negative, i.e., adverse liquidity, productivity, or uncertainty shocks, would be a natural impetus for high asset valuations and low risk premia, as we find. To the contrary, in asset pricing models without an explicit monetary policy, these shocks are a source of high risk premia and low or at best modest asset valuations.

Second, the adverse shocks that may have accompanied the financial crisis and given rise to low natural rates have long since dissipated, yet the real federal funds rate has remained strongly situated in negative territory well after this time, suggesting that the natural rate may have since risen above the policy rate.[10] But such a deviation between the natural rate and the policy rate, were it to exist, would be entirely consistent with the U.S. Central Bank's explicit policy of forward guidance in the post-crisis period, which is a commitment to hold the short rate at or near zero well into the future, beyond the time when it would ordinarily be appropriate according to the parameters of previously estimated systematic policy rules. This interpretation is consistent with our finding that the federal funds rate exhibits more responsiveness to output and less to inflation in the high asset valuation subsamples that include the post-crisis period than in the low valuation subperiod.

Third, estimates of the natural rate vary depending on the model, but often imply values for the natural rate that deviate persistently from the observed real policy rate, even over longer horizons. We obtained recent estimates of the natural rate from Del Negro, Giannone, Giannoni, and Tambalotti (2017) (DGGT) computed from a large-scale, quantitative DSGE

[10] Rajan (2013) questions whether the natural rate is strongly and persistently negative in the post-crisis period.

model and find that, on average, the quarterly real funds rate was 24 basis points above the quarterly natural rate in the low valuation subperiod, while it was 36 basis points *below* in the high valuation subperiods. This suggests that monetary policy was more hawkish in low valuation than in high valuation subperiods. Figure 6 further plots the actual quarterly real funds rate against the DGGT estimate of the 10-year-ahead (forward) natural rate from the DSGE model, and the DGGT model-free estimate of the steady state natural rate from a VAR, both slower-moving or smoothed components of the highly volatile quarterly natural rate estimates. We would expect that over the extended years of our subperiods, deviations between these smoothed series and the quarterly funds rate would average close to zero. But this is not what we find. The figure shows that the funds rate was typically well above the natural rate over most of the low valuation subperiod, but well below during the two high valuation subperiods, and almost entirely so in the post-millennial period.

## 3.4 Asset Market Regimes and Macro Fundamentals

Apart from monetary policy, one might reasonably attribute high asset valuations to more favorable prospects for economic growth and/or lower uncertainty about that growth. Returning to Figure 3, however, we find 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 five and ten years ahead than is the low asset valuation subperiod in the middle of the sample. High asset valuation regimes are also marked by significantly lower expected R&D growth, and weaker investment growth. Figure 7 shows the analogous results for a second MS-VAR, which includes the change in the labor share and stock market dividend growth, in place of investment and R&D growth. The latter figure shows that the high asset valuation subperiods are associated with an expectation of persistent declines in the labor share, whereas the low asset valuation subperiod is characterized by the expectation of a stable labor share. Thus, the boom periods for asset values are associated with broad-based economic weakness and deteriorating payouts to workers. By contrast, there is some evidence that fundamentals for shareholders improve in the form of faster dividend growth, though the credible sets for this result are wide.

In theory, higher asset valuations could be the result of lower expected economic uncertainty (e.g., Lettau, Ludvigson, and Wachter (2008)). But this is also not what we find. Figure 8 reports the median and 68% credible sets of the conditional standard deviation of each variable in the MS-VAR plus the RIR. The conditional standard deviation represents a statistical measure of uncertainty. The result in Figure 8 shows that macroeconomic 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. This is true for uncertainty about GDP growth, inflation, investment growth, and R&D growth. Note that the opposite finding holds for the nominal and real federal funds rate volatility, which is higher in the low asset valuation/high interest rate middle subperiod. This may be consistent with a more active role of the Federal Reserve in stabilizing inflation in the middle subperiod. More aggressive interest rate responses to counter higher inflation and/or a lower output gap may reduce macroeconomic volatility but increase FFR volatility. Of course, this observation provides no support for the hypothesis that the high asset valuation regimes were the product of low economic uncertainty.

Table 3 reports, for the baseline MS-VAR, the means and standard deviations of the real interest rate, GDP growth, R&D growth, and investment growth, conditional on being in a particular regime $i$. Table 4 reports the same output for the second MS-VAR that investigates the change in the labor share and dividend growth. For each draw from the posterior distribution of the MS-VAR parameters, we compute means and standard deviations conditional on being in regime $i$ (see the Appendix for the precise calculation). This procedure gives an entire posterior distribution that we then summarize with the median and 68% posterior credible sets, reported in the Tables. We refer to these as *conditional steady state* values for the moments.

The high asset valuation regime is characterized by sharply lower mean real policy rates and lower uncertainty about those rates, while the opposite is true for the low asset valuation regime. By contrast, Tables 3 and 4 together show that the high asset valuation regime is characterized by lower mean economic growth, lower mean investment growth, lower mean R&D growth, declines in the labor share, faster dividend growth, and higher uncertainty about all of these variables. In contrast to the high valuation subperiods, the low valuation subperiod is marked by increases in the labor share. Differences in the volatility measures across regimes tend to have wide credible sets, however, indicating a large degree of uncertainty over shifts in volatility. But there is no evidence that volatility of fundamentals was significantly lower in the high asset valuation subperiods, as one might expect. Because the conditional steady states do not depend on the estimated transition matrix, these estimates show that the main conclusions are robust to estimation error on the transition matrix.

# 4 Reaching for Yield?

We have found that while persistently low policy rates are associated with persistently high asset valuations, this is not because either signal strong economic growth, favorable changes in inflation, or low uncertainty. This suggests that persistent changes in monetary policy affect asset valuations not because they have long-lasting effects on macroeconomic fundamentals, but 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 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 also fall, further stimulating risky asset valuations. 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]

One way to assess the possibility of RFY is to exploit differences across assets. 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 individual stock market portfolios that exhibit large cross-sectional variation in return premia. An extension of the approach is also used to study the behavior of the risk premium on the aggregate stock market, as explained below.

To motivate the analysis, 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 $\theta_t$. Vuolteenaho (2000) shows that $\theta_t$ 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 $\theta_t$ depends on the present discounted value (PDV) of expected excess returns (risk premia), expected risk-free rates, and expected earnings.

Given two portfolios $x$ and $y$, the spread in their book-market ratios, $\theta_{x,t} - \theta_{y,t}$, is given

[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 genuine covariance with systematic risk factors or mispricing, or both.

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

by:[13]

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

Note that the risk-free rate has no affect on this spread, since all portfolios are affected in the same way by the risk-free rate. Instead only the risk premium differential and expected earnings differential affect the book-market spread. Since RFY 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:

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

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 $i$ with a tilde as

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

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 ($i = 1$) to low ($i = 2$) interest rate regime, the adjusted book-market ratio of $x$ should fall more than $y$, 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:

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

Thus 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. Equivalently, in a switch from a high to low interest rate regime, the PDV of risk premia on the high return premium portfolio should fall more than that of the low return premium portfolio.

## 4.1 Risk Premia Spreads on Equity Portfolios

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 a MS-VAR, now using stock market data rather than macro data. Just as in the previous

[13] This derivation follows Cohen, Polk, and Vuolteenaho (2003) and imposes the approximation that $\rho$ is constant across portfolios. Cohen, Polk, and Vuolteenaho (2003) find that the approximation error generated by this assumption is small.

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. MS-VARs can be used to estimate the difference-in-difference (9) for any two stock portfolios with different average return premia, or it can be used to directly compute the PDV of risk premia $\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t r_{t+1+j}$ for any one asset such the aggregate stock market portfolio. We do both exercises here.

We first present results on how the difference-in-difference has behaved over time and across our regimes using data on individual stock market portfolios that exhibit large cross-sectional variation in return premia. Specifically, we use the equity return data available from Kenneth French's Dartmouth website on portfolios formed from double-sorting all stocks in the AMEX, NYSE, NASDAQ into categories on the basis of five size categories and five BM categories, and alternatively single-sorting into 10 momentum categories based on recent past return performance.[14] We then use CRSP/Compustat to construct the BM ratios of the corresponding portfolios. It is well known that high BM portfolios earn much higher average returns than low BM portfolios, exhibiting the so-called value-spread, especially in the small size quintiles. Along the momentum dimension, recent past winner stocks earn much higher returns than recent past losers. These statistical facts are evident from Table 5, which reports sample statistics for returns on two value spread portfolios–those long in the extreme value portfolio (highest BM ratio) and short in the extreme growth portfolio (lowest BM ratio) while being in the smallest and second smallest size quintile. The same is reported for a momentum spread portfolio–the portfolio that is long in the extreme winner portfolio (M10) and short in the extreme loser portfolio (M1). The annualized Sharpe ratio for the smallest value spread portfolio is 0.62 with an annualized mean return of 10%. Similarly, the momentum spread portfolio has an annualized Sharpe ratio of 0.64 and annualized mean return of over 15%. Both of these strategies have much higher annualized Sharpe ratios and average return premia than the CRSP value-weighted stock market return in excess of the three-month Treasury bill return, where the corresponding numbers are 0.33 and 0.59, respectively.

We estimate a MS-VAR for the former three equity spread-portfolios along with other data that are predictor variables for the returns on these portfolios, chosen on the basis of an Akaike Information Criterion (AIC) selection procedure. The variables included in the MS-VAR are: (a) the momentum return spread, i.e., the difference between the excess return of the extreme

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

winner (M10) portfolio and the excess return of the extreme loser (M1) portfolio; (b) the value return spread (S1), i.e., 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; (c) the value return spread (S2), i.e., the difference between the excess return of the size 2 high BM portfolio and the excess return of the small size 2 low BM portfolio; (d) the momentum BM spread: the difference between the logarithm of the BM ratio of the M10 and M1 portfolios; (e) the value BM spread (S1): 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; (f) the value BM spread (S2): the difference between the logarithm of the BM ratio of the size quintile 2 high book-market portfolio and the logarithm of the BM ratio of the size 2 low book-market portfolio; (g) the real FFR (FFR minus inflation); (h) the excess return on the small value portfolio.[15] The sample for this estimation is 1964:Q1-2013:Q4.[16]

Our objective is exploit the heterogeneity across these portfolios 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 *regime average* values of the adjusted BM *spreads* between the high and low return premia portfolios, $\bar{\tilde{\theta}}^{i}_{xy} \equiv \bar{\tilde{\theta}}^{i}_{x} - \bar{\tilde{\theta}}^{i}_{y}$, for each regime $i$. The regime average value of $\bar{\tilde{\theta}}^{i}_{xy}$ is defined to be the expected value of $\tilde{\theta}_{xy,t}$, conditional on being in regime $i$ today and on the variables of the VAR being equal to their conditional steady state mean values for regime $i$. (The Appendix gives formal expressions for the regime average, and explains how they are computed from the MS-VAR parameters.) For each draw of the VAR parameters from the posterior distribution of these parameters, we compute the median and 68% credible sets for $\bar{\tilde{\theta}}^{i}_{xy}$, which are reported in Table 6. The high ($x$) and low ($y$) 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). The third row reports the analogous values for the regime average of 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 $\left(\tilde{\theta}^{1}_{x,t} - \tilde{\theta}^{1}_{y,t}\right) - \left(\tilde{\theta}^{2}_{x,t} - \tilde{\theta}^{2}_{y,t}\right)$, as

[15] The BM spreads are included because they represent the natural valuation ratios for the portfolio return spreads that we are trying to predict (Cohen, Polk, and Vuolteenaho (2003)). The real FFR and the excess return on the small value portfolio are selected based on the Akaike Information Criterion (AIC) among a set of possible additional regressors. The five Fama/French factors (Fama and French (2015)) are considered as possible additional regressors, but not selected based on the AIC. We note that our value spread returns behave similarly to the HML factor of Fama and Fench. The Online Appendix provides additional details on the variable selection procedure.

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.

[16] 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.

implied by the VAR estimates. The last row reports the posterior probability that risk premia decline in low interest rate regime, computed as the percentage of draws from the posterior distribution of regime averages for which risk premia are lower in regime 2 than in regime 1. 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.

Table 6 shows that the adjusted BM spreads $\overline{\overline{\theta}}^{i}_{xy}$ between the high and low return premia portfolios are positive in both regime 1 and regime 2. This is not surprising because portfolios that have higher risk premia should have lower market values relative to their book values, holding fixed expected earnings. As equation (8) shows, this simply indicates that the ostensively riskier portfolio has a higher PDV return premium on average, regardless of the regime. 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, in a shift from high to low interest rate regime, the adjusted book-market ratios of high return/high risk portfolios fall more 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 interest rates than do less risky/lower return assets.

The third row of Table 6 also reports the 68% posterior credible intervals in parentheses for the difference-in-difference. For the BM spreads, although the median value for the diff-in-diff is always positive, the 68% posterior credible intervals include negative values. This is not true of the momentum spread W-L or the excess market return (discussed below) where the credible sets are narrow and include only positive values. However, because the distribution of the diff-in-diff displays substantial negative skewness, the posterior probability assigned to a decline in premia during low interest rate regimes is nonetheless quite high in all cases: 88%, 64%, and 70%, for the Momentum spread, the S1 BM spread, and the S2 BM spread, respectively.

These results are supportive of a channel that implies an increased appetite for risk-taking in low interest rate environments. The case of momentum in particular deserves emphasis. The results indicate that the spread in adjusted BM spread 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. Momentum investing is known to be among the most volatile equity investment strategies studied, one that is subject to infrequent but extreme crashes in times of crisis or panic (Daniel and Moskowitz (2013)).

We also estimate the PDV of risk premia as it evolves over the sample, rather than an average value conditional on a regime. These values are estimated as the time $t$ VAR forecasts, i.e. conditional expected values, of the PDV on the right-hand-side of (8).[17] Given the posterior

[17]The expectation $\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot|\mathbb{I}_t)$, computed as above where $\mathbb{I}_t$ includes knowledge of the regime in place at

distribution of the VAR parameters, these forecasts have their own posterior distribution. Figure 9 reports the median values of these forecasts as solid (blue) lines, while the regime averages are indicated by the dashed (red) lines. Although the risk premia are volatile, it is clear that they fluctuate around a lower value in the high asset valuation/low interest rate regime than they do in the low asset valuation/high interest rate regime. The portfolio strategies 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. Estimated risk premia then return to low levels in the post-crisis ZLB period.

## 4.2 The Risk Premium on the Aggregate Stock Market

We follow a similar procedure to investigate how the risk premium on the overall stock market has evolved over the sample. To do so, we form an estimate of $\sum_{j=0}^{\infty} \rho^j \mathbb{E}_t r_{t+1+j}$ for the aggregate stock market from a MS-VAR specification that includes the following variables: (a) the market excess return, computed as the difference in the CRSP value-weighted stock market return (including dividend redistributions) and the three-month Treasury bill rate; (b) the inverse valuation ratio $-cay^{MS}$; (c) the small stock value spread (log-difference in the book to market ratio of the S1 value and S1 growth portfolio); (d) the SMB factor from Fama and French; (e) the HML factor from Fama and French. The inverse valuation ratio $-cay^{MS}$ is included because $\epsilon^{MS}$ strongly predicts future stock market returns (see the next section).[18] The small stock value spread and the SMB and HML Fama-French factors are included because they improve the AIC criterion.[19]

The estimated evolution of the PDV of the market risk premium is displayed in the first panel of Figure 9. It is substantially lower in the high asset valuation/low interest rate regime than in the low asset valuation/high interest rate regime, and reaches lows or near-lows in the post-2000 period and the post-crisis period with interest rates in the ZLB range. The premium exhibits a sharp increase during the years corresponding to the financial crisis, but it declines again subsequent to the crisis.

Returning to Table 6, the first column reports the regime average value for the PDV of the

time $t$, $Z^t$, and the VAR parameters for each regime.

[18] Recall that $cay^{MS}$, unlike $\epsilon^{MS}$, includes the Markov-switching constant, which provides information on the low frequency swings in the risk premium. Accounting for these very low frequency swings was less important for the previous short-horizon forecasting exercise. But it is potentially valuable for estimating the infinite PDV of expected future return premia when used in a predictive VAR that takes into account the probability of eventually switching out of a regime.

[19] Campbell and Vuolteenaho (2004) first reported the forecasting power of the small stock value spread for the market return. The specification of this VAR differs from that used above to estimate adjusted BM spreads across portfolios, for two reasons. First, in practice the two sets of variables call for different predictors. Second, specifying a single VAR for both the spreads and the market premium would require including a large number of variables with a consequent significant loss in terms of precision of the estimates.

market premium across the two regimes (first and second row,) the distribution of the difference between the premia across regimes (third row,) and the probability that the PDV of the market risk premium falls in the low interest rate regime (last row.) The PDV of the expected market excess return is estimated to be positive in both regimes, but is substantially larger under the low valuation/high interest rate regime (around 1.5 times larger). This difference is statistically significant, as suggested by the credible sets reported in the third row. Accordingly, the last row shows that the posterior probability that premia decline when moving to the high valuation regime is large (85%.)

In summary, the results in this section suggest that low interest rate regimes are associated with low risk premia in equity markets, supportive of reaching for yield theories. The findings underscore the challenges for theories that explain persistently low interest rate environments with shocks that shift in the composition of wealth toward more risk averse or more pessimistic investors (e.g., Barro and Mollerus (2014); Caballero and Farhi (2014); Hall (2016)), which under standard calibrations imply that low interest rates coincide with higher rather than lower risk premia, contrary to what we find. The findings above suggest that, not only are risk premia lower conditional on being in a low interest rate regime, but the estimated historical variation in these premia reach lows or near-lows in the post-millennial period and again at the onset of the ZLB period.

# 5 Sources of Variation in $cay_t^{MS}$

## 5.1 Forecasts of Excess Stock Market Returns

The above sections focused on the low frequency regimes in $cay^{MS}$, which were found to be strongly correlated with interest rates. The variable $cay$ has traditionally been used as a stock market return 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 excess returns, rather than higher values for $c$ and/or $y$. This section presents evidence that adjusting for the persistent interest rate regimes documented here allows $cay^{MS}$ to forecast excess stock returns better than the traditional $cay$, especially over short- to medium-horizons.

Before presenting these results, its instructive to consider why adjusting for persistent regimes is likely to improve forecasting power. In the fixed coefficient version of $cay$, where there is no adjustment for regimes, there is a one-to-one mapping from $cay_t^{FC}$ to the cointegrating residual $\epsilon_t^{FC}$. By contrast, $cay_t^{MS}$ is comprised of two components, the Markov-switching constant $\alpha_{\xi_t}$, which does adjust for regimes, and the cointegrating residual $\epsilon_t^{MS}$. Unlike $cay_t^{FC}$, $\epsilon_t^{MS}$ takes into account that asset valuations can be persistently high or persistently low. This has the potential to improve forecasts. Knowing that we are in a period of high (low) val-

uation allows us to adjust forecasts appropriately, so that we do not expect asset valuations to immediately revert to the low (high) valuation regime, but rather to fluctuate around the persistently varying mean $\alpha_{\xi_t^\alpha}$ for a while. Because $\epsilon_t^{MS}$ measures fluctuations in $cay_t^{MS}$ around $\alpha_{\xi_t^\alpha}$, it explicitly recognizes the persistence of regimes so that forecasts of stock returns over relatively short horizons (say, less than four years) are improved. Of course, as the horizon increases, the possibility of a regime change becomes more likely.

Table 7 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. The table compares the forecasting power of $cay_t^{FC}$, and $\epsilon_t^{MSfilt}$ and $\epsilon_t^{MS}$, the high frequency components of $cay_t^{MS}$ based on filtered and smoothed operabilities, respectively. The top panel reports full sample forecasts. The bottom panel reports the results of forecasts based on fully recursive estimates of these measures using data only up to time $t$, denoted $cay^{FCrec}$ and $\epsilon^{MSrec}$, respectively.[20] The recursive estimates use no forward looking data to estimate any of the parameters, including the regime probabilities, regimes values, or transition probabilities. We report the coefficient estimates on the regressor, the Newey and West (1987) corrected $t$-statistic, and the adjusted $R^2$ statistic.

Our use of the recursive estimates of $\epsilon^{MS}$ underscores a distinction with other "structural change" methodologies that rely on ex-post structural breaks capable of being identified only with hindsight (e.g., Lettau and Van Nieuwerburgh (2008)). The methodology used here, which specifies recurrent regime changes with probabilistic switches, requires only that the researcher consider the possibility of regime changes in order to estimate the appropriate model. Once this is done, the probability of a switch in the future, along with the rest of the model parameters, can be estimated in real time using information only up to time $t$, allowing the researcher to estimate $cay_t^{MS}$ and its two components with data only available at time $t$.

The top panel of 7 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 $\epsilon_t^{MSfilt}$ and $\epsilon_t^{MS}$ than they are for $cay_t^{FC}$. The comparison is more stark if we focus on recursively estimated values of $cay$ to full sample values. 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 $\epsilon_t^{MSrec}$ explains 42%. Moreover, in this sub-sample, $cay_t^{FC}$ has little forecasting power for excess returns at all but the longest horizon, whereas $\epsilon_t^{MSrec}$ has much stronger forecasting power.

[20]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. These variables are then used to forecast returns over the entire subsample from 1981:Q1-2013:Q3.

Table 7 shows that $cay_t^{FCrec}$ also has stronger predictive power than $cay_t^{FC}$ over this subsample. This happens because, by recursively estimating the parameters in $cay_t^{FC}$, we allow them to change *every period*. For this reason, a recursively estimated fixed-coefficient model can "compete" with the Markov-switching model, which explicitly models shifts in the parameter. This finding hardly provides support for the hypothesis that the fixed-coefficient model is a better description of the data than the Markov-switching model. (Indeed, this hypothesis is explicitly rejected by a comparison based on the *BIC* criterion of these models, as documented in the Appendix.) 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 $cay_t^{FC}$.

Table 8 reports mean-square forecast errors (MSEs) from out-of-sample forecasts.[21] As for the fully recursive estimates, these parameters are estimated in real time and do not rely on estimates of break dates or parameters across regimes that could not have been apparent to investors in real time. All versions of *cay* also have lower lower MSEs than a simple autoregressive forecasting model or a model that uses only the (constant) sample mean of excess returns as a predictor. Among those versions that are estimated using the full sample, the two Markov-switching cointegrating residuals, $\epsilon_t^{MSfilt}$, and $\epsilon_t^{MS}$, are 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 $\epsilon_t^{MSrec}$ 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 reason, it is notable that $\epsilon_t^{MSrec}$ preforms as well (and slightly better at long horizons) as $cay_t^{FCrec}$, given that the former has more parameters that require estimation over short subsamples of our quarterly dataset. Postwar samples of the size currently available are 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 $cay_t^{MS}$ and of its decomposition into $\alpha_{\xi_t}$ and $\epsilon_t^{MS}$.

[21] 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.

## 5.2 A Variance Decomposition of $cay_t^{MS}$

With $cay_t^{MS} = \alpha_{\xi_{t+h}^{\alpha}} + \epsilon_t^{MS}$, it is straightforward to compute the relative importance of the two sources of variation:

$$\mathbb{V}_t\left[cay_{t+h}^{MS}\right] = \mathbb{V}_t\left[\alpha_{\xi_{t+h}^{\alpha}}\right] + \mathbb{V}_t\left[\epsilon_{t+h}^{MS}\right] \tag{10}$$

where $\mathbb{V}_t\left[x_{t+h}\right]$ is the variance of $x_{t+h}$ conditional on the information available at time $t$. In deriving the analytical formula above, we have taken advantage of the assumption that $\xi_t^{\alpha}$ and $\epsilon_t^{MS}$ are estimated as two uncorrelated disturbances. Since $\mathbb{V}_t\left[\cdot\right]$ depends on information available at time $t$, we condition on being in a certain regime today, but employ a decomposition that takes into account the possibility of switching out of that regime in the future. As the horizon increases, the conditional variance terms converge to their unconditional counterparts.

The conditional variance of the Markov-switching constant can be computed based on the parameter estimates for $\alpha_{\xi_t^{\alpha}}$ and the transition matrix $H^{\alpha}$ using the formulas for the conditional variance of a Markov-switching process presented in Appendix. Our estimation does not provide a corresponding parametric specification for the cointegrating residual $\epsilon_t^{MS}$. Thus, we fit an AR(1) model for $\epsilon_t^{MS}$:

$$\epsilon_t^{MS} = \underbrace{\rho_{\epsilon}}_{0.8186}\epsilon_t^{MS} + \underbrace{\sigma_{\epsilon}}_{0.0069}u_t,\ u_t \sim N\left(0,1\right),$$

and then use the parameter estimates of the AR(1) process to compute the conditional variance of $\epsilon_{t+h}^{MS}$. The sum of the two components give us the conditional variance of $cay_{t+h}^{MS}$.

Table 9 reports the results from the decomposition of the left-hand-side of (10) into its two components on the right-hand-side, at the posterior mode estimates of all parameters. The upper panel reports the variance decomposition conditional on being in the low valuation regime at time $t$, while the lower panel reports results conditional on being in the high valuation regime at time $t$. In each panel, the first row reports the contribution of $\epsilon_t^{MS}$, whereas the second row reports the contribution of $\alpha_{\xi_{t+h}^{\alpha}}$.

A sizeable fraction of the expected volatility in $cay_t^{MS}$ can be attributed to $\epsilon_t^{MS}$. This is especially true when looking out over shorter horizons, since the high persistence of the intercept regimes implies that the expectation of staying in a given regime is quite high. For example, conditional on being in the low (high) valuation regime at $t$, $\epsilon_t^{MS}$ accounts for 77% (81%) of the $h = 1$ quarter variation in $cay_t^{MS}$ and 68% (73%) of the $h = 4$ quarter variation. As the horizon increases, regime changes in $\alpha_{\xi_{t+h}^{\alpha}}$ become more important. This is not surprising, given that regime changes are quite persistent. However, even at a 20-year horizon, $\epsilon_t^{MS}$ still explains 33% to 37% of the volatility of $cay_t^{MS}$. As the horizon increases, the conditional variances converge to their unconditional counterparts. At the $h = \infty$ horizon, the importance of fluctuations in $\epsilon_t^{MS}$ is close to 30%.

Putting all the previous results together, it is tempting to conclude that shifts in the inter-

cept term $\alpha_{\xi^{\alpha}_{t+h}}$ are attributable solely to interest rate movements while shifts in the residual term $\epsilon^{MS}_t$ are attributable to risk premia variation. However, nothing in the variance decomposition prevents the residual term from having a component correlated with interest rates (at frequencies higher than that captured by the intercept term), or conversely prevents the intercept term from having a component correlated with risk premia (at frequencies lower than that captured by the residual term). Indeed the latter possibility is supported by the evidence in the previous section, which is consistent with reaching for yield across the interest rate regimes. Taken together, the findings suggest that variation in $cay^{MS}_t$ is driven at higher frequencies by $\epsilon^{MS}_t$, which is associated with forecastable variation in excess returns, while its variation at lower frequencies, driven by $\alpha_{\xi^{\alpha}_{t+h}}$, is correlated with both interest rates and expected return premia. In this respect, our results for $\alpha_{\xi^{\alpha}_{t+h}}$ complement evidence in Lustig, Van Nieuwerburgh, and Verdelhan (2013), who find that their measure of the ratio of aggregate (human plus non-human) wealth to consumption, varies with bond market rates. Our findings differ from these authors in that $cay^{MS}_t$ also contains a component attributable to fluctuations in equity market risk premia, both because movements in $\alpha_{\xi^{\alpha}_{t+h}}$ are associated with movements in risk-premia at low frequencies, and because movements in $\epsilon^{MS}_t$ are associated with risk premia movements at higher frequencies.

# 6 Other Valuation Ratios

This section briefly comments on the behavior of other stock market valuation ratios over our sample. Many such ratios, e.g., price-dividend ratios for the aggregate stock market, or price-payout ratios, appear to differ from *cay* in that they exhibit trends. For example, price-dividend ratios drift up over the post-war period. The methods employed here are not suited to explaining non-stationary trends. For the purposes of this paper, we address this crudely by simply removing a trend in these ratios before analyzing the data. Rather than removing only the lowest frequency components, however, we use a band-pass filter to examine frequencies that correspond to cycles between 10 and 50 years, or "medium-term" components. These frequencies roughly coincide with the persistence of our previously estimated regimes for $cay^{MS}$ and so form a natural basis for comparison with those results.

Figure 10 plots the medium-term components of five different stock market valuation ratios, overlaying the $cay^{MS}$ regime subperiods on the figure, with the low *cay*-valuation subperiod indicated in gray shading and the high valuation subperiods at the beginning and end of our sample indicated in white shading. The five valuation ratios plotted are a CRSP value-weighted stock price-dividend ratio on a portfolio that does not reinvest dividends, the Flow of Funds (FOF) price-payout ratio, the FOF price-dividend ratio, Shiller's price-earnings ratio[22], and

[22] http://www.multpl.com/shiller-pe/

the value-weighted price-dividend ratio for all firms in NYSE, NASDAQ, and AMEX from a COMPUSTAT/CRSP merge. The figure shows that the medium-term components of these five other valuation ratios exhibit similar high-low-high valuation patterns over the same subperiods that characterize the high-low-high $cay_t^{MS}$ valuation regimes. Moreover, the average values of these series in each regime differ noticeably, especially for the price-dividend and price-payout ratios.

# 7 Conclusion

This paper documents a novel set of statistical facts on how asset values, macroeconomic aggregates, short-term interest rates, and risk premia have comoved over longer periods of time in post-war U.S. data. We present evidence of infrequent shifts, or regimes, 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. Evidence from a Markov-Switching VAR shows that these low frequency swings in post-war asset valuation are strongly associated with low frequency mean shifts in the Federal Reserve's primary policy interest rate, with low VAR forecasts for the real federal funds rate associated with high asset valuations, and vice versa. High asset valuation regimes are also characterized by evidence of less monetary policy activism toward inflation and greater activism toward output growth.

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 rational optimism about the future in the form of expectations for stronger long-run economic growth or lower uncertainty about that growth. Indeed, high valuation regimes, including the post-millennial period, are marked by economic weakness in GDP growth, investment growth, and R&D growth, along with declines in the labor compensation share of GDP. The one exception to this evidence of declining expected prosperity in high valuation regimes is the stock market itself, where there is some evidence that the fundamentals for shareholders improve.

Finally, we present evidence that the high valuation/low interest rate regimes are marked by significantly lower equity market risk premia, consistent with the hypothesis that investors' willingness to tolerate risk in equity markets rises when the long-run expected value of the real federal funds rate is low. 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.

# 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.

Baele, L., G. Bekaert, S. Cho, K. Inghelbrecht, and A. Moreno (2015): "Macroeconomic regimes," *Journal of Monetary Economics*, 70, 51–71.

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

Bekaert, G., M. Hoerova, and M. L. Duca (2013): "Risk, Uncertainty and Monetary Policy," *Journal of Monetary Economics*, 60(7), 771–88.

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.

Bigio, S. (2015): "Endogenous liquidity and the business cycle," *The American Economic Review*, 105(6), 1883–1927.

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 T. Vuolteenaho (2004): "Bad Beta, Good Beta," *American Economic Review*, 94(5), 1249–1275.

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.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

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.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003): "The value spread," *The Journal of Finance*, 58(2), 609–641.

Coimbra, N., and H. Rey (2017): "Financial cycles with heterogeneous intermediaries," Discussion paper, National Bureau of Economic Research.

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.

Costa, O., M. Fragoso, and R. Marques (2004): *Discrete-Time Markov Jump Linear Systems*. Springer, New York.

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

Del Negro, M., G. Eggertsson, A. Ferrero, and N. Kiyotaki (2017): "The great escape? A quantitative evaluation of the Fed's liquidity facilities," *The American Economic Review*, 107(3), 824–857.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017): "Safety, liquidity, and the natural rate of interest," Available at SSRN: https://ssrn.com/abstract=2967235.

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

Diamond, D. W., and R. G. Rajan (2012): "Illiquid banks, financial stability, and interest rate policy," *Journal of Political Economy*, 120(3), 552–591.

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

Eggertsson, G., and M. Woodford (2003): "Zero bound on interest rates and optimal monetary policy," *Brookings papers on economic activity*, 2003(1), 139–233.

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

FARHI, E., AND J. TIROLE (2012): "Collective moral hazard, maturity mismatch, and systemic bailouts," *The American Economic Review*, 102(1), 60–93.

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

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.

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

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

KLIEM, M., L. MEYER-GOHDE, AND S. MOYEN (2017): "(Un)Expected Monetary Policy Shocks and Term Premia: A Bayesian Estimated Macro-Finance Model," https://www.wiso.uni-hamburg.de/fachbereich-vwl/professuren/meyer-gohde/bilder/unexpected-monetary-shocks.pdf.

KURLAT, P. (2013): "Lemons markets and the transmission of aggregate shocks," *The American Economic Review*, 103(4), 1463–1489.

LARRAIN, B., AND M. YOGO (2008): "Does firm value move too much to be justified by subsequent changes in cash flow?," *Journal of Financial Economics*, 87(1), 200–226.

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

——— (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., 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.

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.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2013): "The wealth-consumption ratio," *Review of Asset Pricing Studies*, 3(1), 38–94.

Miranda-Agrippino, S., and H. Rey (2015): "US monetary policy and the global financial cycle," *NBER working paper*, 21722.

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.

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.

Piazzesi, M., and M. Schneider (2015): "Payments, credit and asset prices," Unpublished manuscript, Stanford University.

Primiceri, G. E. (2005): "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, 72(3), 821–852.

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.

Rossi, B., and T. Sekhposyan (2016): "Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts," *Journal of Applied Econometrics*, 31(3), 507–532.

Rudebusch, G. D., and E. T. Swanson (2012): "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *American Economic Journal: Macroeconomics*, pp. 105–143.

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

Stein, J. C. (2013): "Overheating in Credit Markets: Origins, Measurement, and Policy Responses," *Lecture at "Restoring Household Financial Stability after the Great Recession: Why Household Balance Sheets Matter", Federal Reserve Bank of St. Louis, St. Louis, Missouri.*

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

Taylor, J. B. (1993): "Discretion versus policy rules in practice," in *Carnegie-Rochester Conference Series on Public Policy*, vol. 39, pp. 195–214.

——— (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.

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

**Parameter Estimates:** $cay^{MS}$

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

**Parameter Estimates:** $cay^{FC}$

| $\alpha$ | $\gamma_a$ | $\gamma_y$ |
|---|---|---|
| 0.8706 | 0.1246 | 0.7815 |
| (0.0345) | (0.0150) | (0.0168) |

**Table 1:** Parameter estimates. The top panel reports 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 lower panel reports 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.

**Long term responses of the FFR**

| | Low Val. regime | High Val. regime | Difference |
|---|---|---|---|
| Inflation | 2.0048 | 1.1379 | 0.8750 |
| | (1.6016,2.8005) | (0.9317,1.3437) | (0.4117,1.7192) |
| Output growth | 1.0178 | 1.6423 | −0.6972 |
| | (−0.0079,2.2248) | (1.1801,2.1968) | (−1.8892,0.6878) |

**Table 2:** Long-term responses to a permanent 1% increase in inflation or GDP growth. This table reports the median and (in parentheses) the 68% posterior credible sets of the long term response of the FFR to inflation and GDP growth conditional on being in a certain regime. The last column reports the distribution for the difference in the long term responses between the low and high asset valuation regimes. The sample spans the period 1955:Q3-2013:Q3.

**Summary statistics for macroeconomic variables**

| | Real Interest Rate | | | GDP growth | | |
|---|---|---|---|---|---|---|
| | Low Val. | High Val. | Diff. | Low Val. | High Val. | Diff. |
| Mean | 3.6729 (2.9927,4.4366) | 0.5638 (0.0130,1.1454) | 3.1015 (2.1029,4.2184) | 2.2999 (1.6115,3.1994) | 1.4221 (0.8696,1.9632) | 0.9271 (−0.1140,2.1806) |
| SD | 2.4376 (2.0175,3.1353) | 1.8229 (1.5487,2.3100) | 0.5890 (−0.0194,1.3029) | 2.3076 (1.9831,2.8225) | 2.8846 (2.5168,3.4306) | −0.5746 (−1.1672,0.0702) |
| | R&D growth | | | Investment growth | | |
| | Low Val. | High Val. | Diff. | Low Val. | High Val. | Diff. |
| Mean | 5.5476 (4.6201,6.7832) | 2.9990 (1.7074,4.2863) | 2.6316 (0.7845,4.7214) | 2.2859 (0.0796,5.0458) | 0.9770 (−0.8985,2.8737) | 1.4307 (−1.8616,5.1755) |
| SD | 4.0985 (3.4764,5.0801) | 5.0998 (4.3961,6.2689) | −0.9623 (−2.2973,0.1982) | 8.6720 (7.5000,10.4987) | 11.1561 (9.8214,13.0066) | −2.4246 (−4.5976,−0.2669) |

**Table 3:** Conditional Steady States. This table reports the median and (in parentheses) 68% posterior credible sets of the conditional mean and standard deviation for the real interest rate, GDP growth, R&D growth, and investment growth based on the VAR estimates conditional on staying in each regime.The sample spans the period 1955:Q3-2013:Q3.

**Summary statistics for dividend growth and change in labor share**

| | Dividend growth | | | Change in labor share | | |
|---|---|---|---|---|---|---|
| | Low Val. | High Val. | Diff. | Low Val. | High Val. | Diff. |
| Mean | 1.5739 (0.4942,2.5969) | 2.8996 (0.3542,5.2595) | −1.3408 (−4.2040,1.6341) | 0.0329 (−0.0658,0.1517) | −0.1486 (−0.2613,−0.0530) | 0.1905 (0.0378,0.3672) |
| SD | 6.0888 (5.5313,6.8447) | 11.2397 (10.0984,12.8371) | −5.0828 (−6.8857,−3.7131) | 0.5433 (0.4669,0.6590) | 0.5851 (0.5112,0.6972) | −0.0427 (−0.1773,0.0935) |
| | Real Interest Rate | | | GDP growth | | |
| | Low Val. | High Val. | Diff. | Low Val. | High Val. | Diff. |
| Mean | 3.4943 (2.8508,4.1097) | 0.5483 (0.0534,1.0554) | 2.9438 (2.0050,3.8169) | 2.3679 (1.7477,3.2405) | 1.3870 (0.8918,1.8921) | 1.0310 (0.1498,2.1705) |
| SD | 2.3287 (1.9513,2.9377) | 1.7667 (1.5048,2.1925) | 0.5594 (−0.0185,1.1956) | 2.2431 (1.9302,2.7608) | 2.8104 (2.4868,3.3197) | −0.5584 (−1.1388,0.0508) |

**Table 4:** Conditional Steady States. This table reports the median and (in parentheses) 68% posterior credible sets of the conditional mean and standard deviation for dividend growth, change in the labor share, the real interest rate, and GDP growth based on the VAR estimates conditional on staying in each regime. The sample spans the period 1955:Q3-2013:Q3.

Annualized Sharpe Ratios and Mean Returns

| Portfolio | SR | Mean | Portfolio | SR | Mean |
|---|---|---|---|---|---|
| Market | 0.3335 | 0.0585 | V-G (S1) | 0.6225 | 0.1047 |
| W-L | 0.6446 | 0.1588 | V-G (S2) | 0.3807 | 0.0637 |

**Table 5:** The table reports annualized Sharpe ratios, "SR," and mean returns, "Mean," for the stock market and 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, S2= second smallest. 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). 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 Market Premium and Book-Market Ratio Spreads**

| | Market | W-L | Val-Gr (S1) | Val-Gr (S2) |
|---|---|---|---|---|
| Regime 1 | 1.6722 (0.9171,2.4646) | 4.2409 (2.8893,5.5322) | 2.5508 (1.4115,3.6787) | 1.5403 (0.4540,2.6152) |
| Regime 2 | 1.0920 (0.3629,1.8055) | 2.9978 (0.9953,4.5221) | 2.2642 (1.2361,3.1899) | 1.1466, (0.0289,2.1074) |
| Diff-in-Diff | 0.5331 (0.0270,1.1670) | 1.1962 (0.2072,2.7763) | 0.2677 (−0.5071,1.2286) | 0.3772, (−0.3803,1.4059) |
| Prob. decline | 0.85 | 0.88 | 0.64 | 0.70 |

**Table 6:** The first two rows report the conditional steady states for the present discounted value of market excess returns and the spread in adjusted (for expected earnings) book-market ratios between the high and low return premia portfolios 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-in-Diff" reports the difference between these spreads across the two wealth ratio/interest rate regimes. The numbers in each cell are the median values of the statistic from the posterior distribution while in parentheses we report 68% posterior credible sets. The last row reports the probability that premia decline when moving from the low valuation regime to the high valuation regime. These probabilities are obtained by computing the fraction of draws from the posterior distribution for which the premia under the high valuation regime are lower than the premia under the low valuation regime.

**Long Horizon Forecasting Regressions: Stock Returns**

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

| | Horizon $h$ (in quarters) | | | | |
|---|---|---|---|---|---|
| $z_t =$ | 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] |
| $\epsilon_t^{MSfilt} = cay^{MS} - \overline{\alpha}_t^{filt}$ | **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] |
| $\epsilon_t^{MS} = cay^{MS} - \overline{\alpha}_t$ | **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] |
| $\epsilon_t^{MSrec} = cay^{MSrec} - \overline{\alpha}_t^{rec}$ | 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 7:** 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; $\epsilon_t^{MSfilt}$ denotes the high frequency component of $cay^{MS}$ using filtered probabilities and $\epsilon_t^{MS}$ denotes the high frequency component of $cay^{MS}$ 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 |
| $\epsilon^{MSfilt}$ | 0.70 | 2.47 | 2.64 | 3.01 | 3.72 |
| $\epsilon^{MS}$ | 0.70 | 2.35 | 2.53 | 2.92 | 3.68 |
| $\epsilon^{FCrec}$ | 0.72 | 2.87 | 4.38 | 5.72 | 6.61 |
| $\epsilon^{MSrec}$ | 0.71 | 2.86 | 4.49 | 5.75 | 6.14 |

**Table 8:** 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, $\epsilon_t^{MSfilt}$ and $\epsilon_t^{MS}$ are the high frequency components of $cay^{MS}$ obtained using filtered and smoothed probabilities, respectively, $cay^{FCrec}$ is the recursively estimated $cay$ with fixed coefficients, and $\epsilon_t^{MSrec}$ is the high frequency component of 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.

**Variance Decomposition for $\mathbf{cay^{MS}}$**

| | 1Q | 1Y | 5Y | 10Y | 20Y | 40Y | Uncond. |
|---|---|---|---|---|---|---|---|
| Regime 1: $\epsilon^{MS}$ | 0.7720 | 0.6842 | 0.5834 | 0.4185 | 0.3305 | 0.2923 | 0.2924 |
| Regime 1: $\alpha_{\xi_t^\alpha}$ | 0.2280 | 0.3158 | 0.4166 | 0.5815 | 0.6695 | 0.7077 | 0.7076 |
| Regime 2: $\epsilon^{MS}$ | 0.8170 | 0.7393 | 0.6448 | 0.4754 | 0.3726 | 0.3156 | 0.2924 |
| Regime 2: $\alpha_{\xi_t^\alpha}$ | 0.1830 | 0.2607 | 0.3552 | 0.5246 | 0.6274 | 0.6844 | 0.7076 |

**Table 9:** This table reports the relative contribution of regime changes and cointegrating residual to the variance of $cay_t^{MS}$. The uppertwo rows report the decomposition conditional on being in the low valuation regime today, while the last two rows report the decomposition conditional on being in the high valuation regime today. While this conditions on being in a certain regime today, the variance decomposition considers the possibility of future regime changes. The sample spans the period 1952:Q1-2013:Q3.

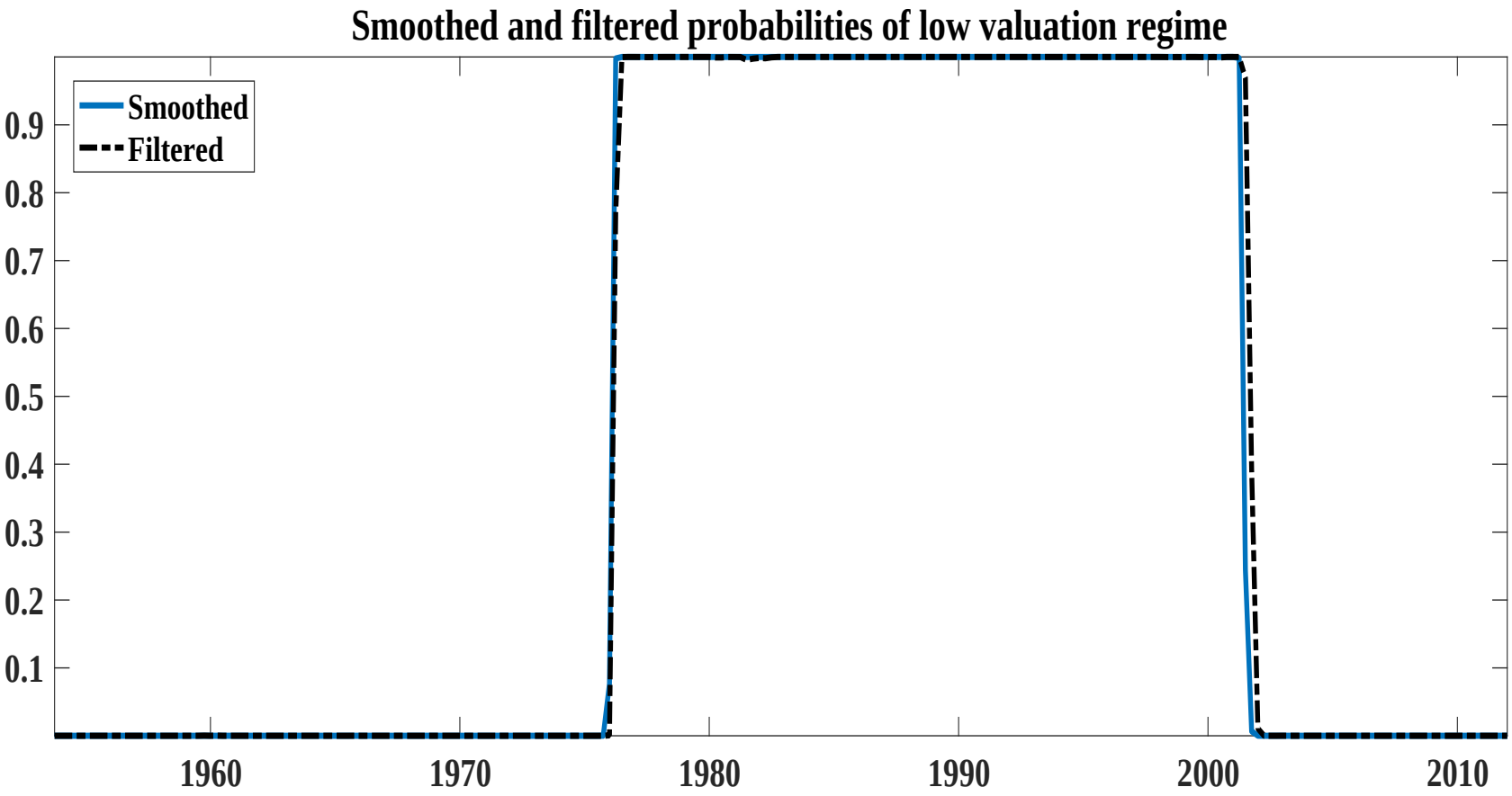



**Figure 1:** Smoothed probability of a low asset valuation regime for the Markov-switching cointegrating relation. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

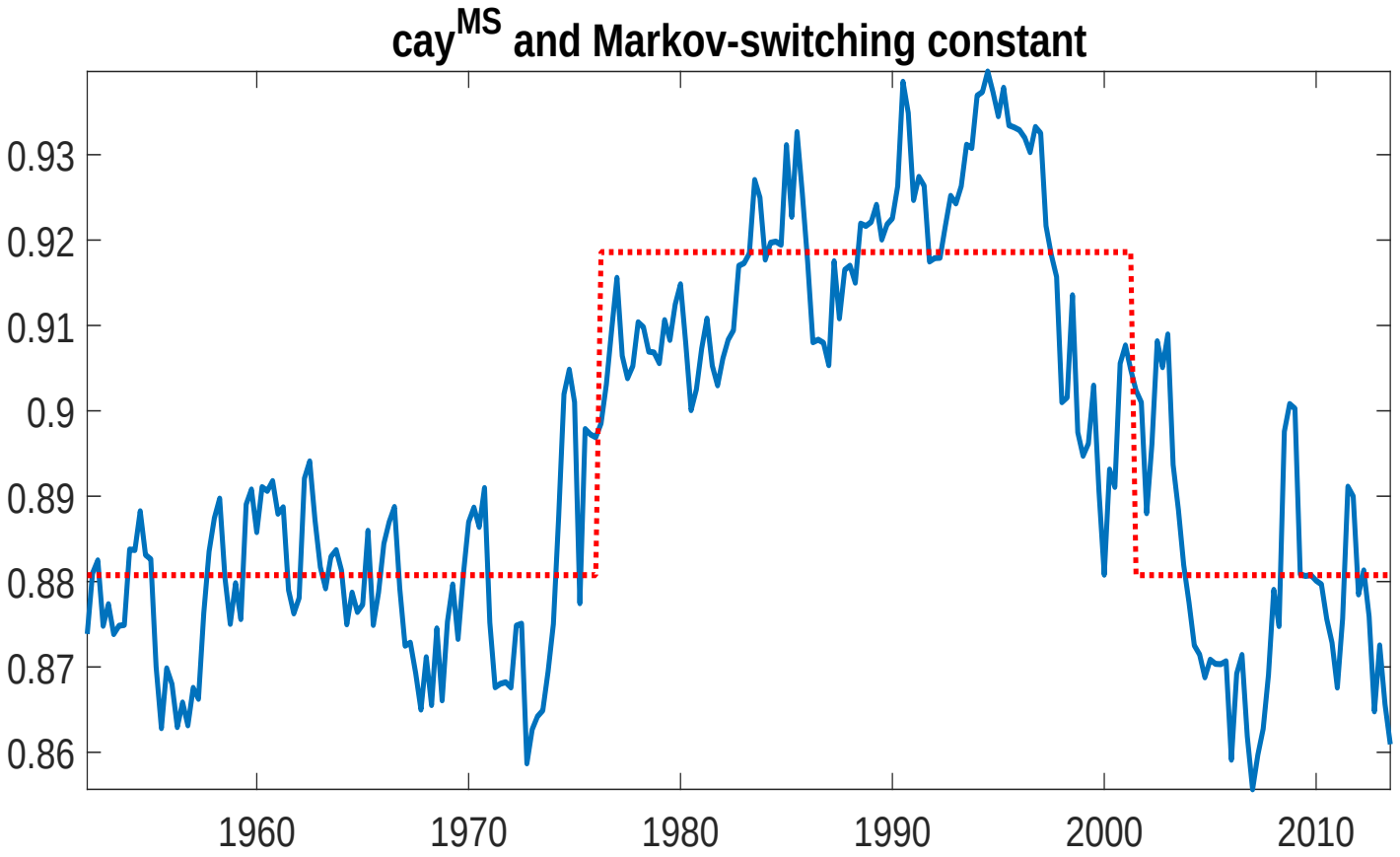



**Figure 2:** The solid line reports the Markov-switching estimated $cay^{MS}$. The red dashed lines is the estimate $\overline{\alpha}_t$. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

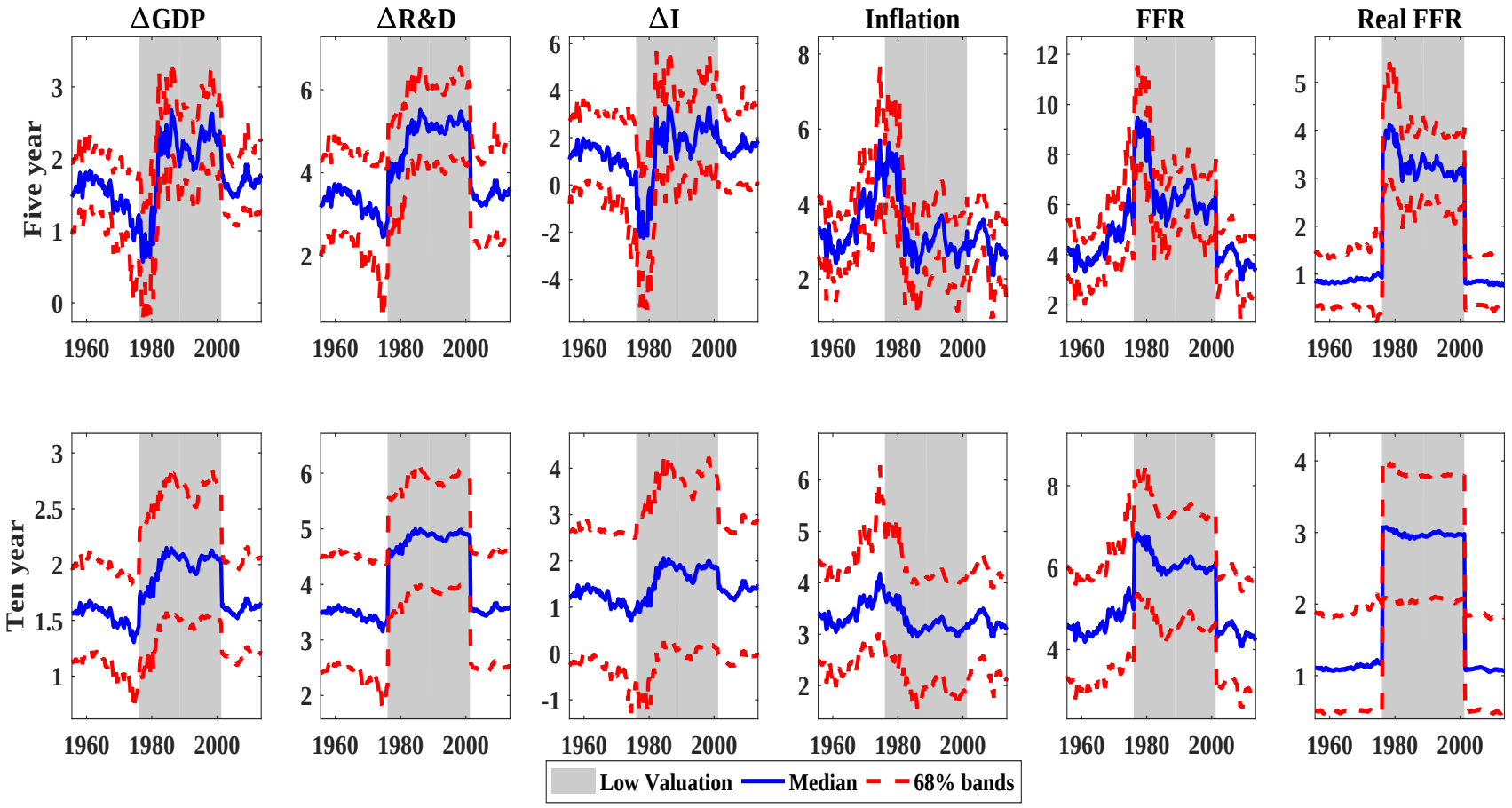



**Figure 3: Conditional expectations from baseline MS-VAR.** The figure reports the conditional expectations based on the baseline MS-VAR at different horizons taking into account the possibility of regime changes. The sample spans 1955:Q3 to 2013:Q3.

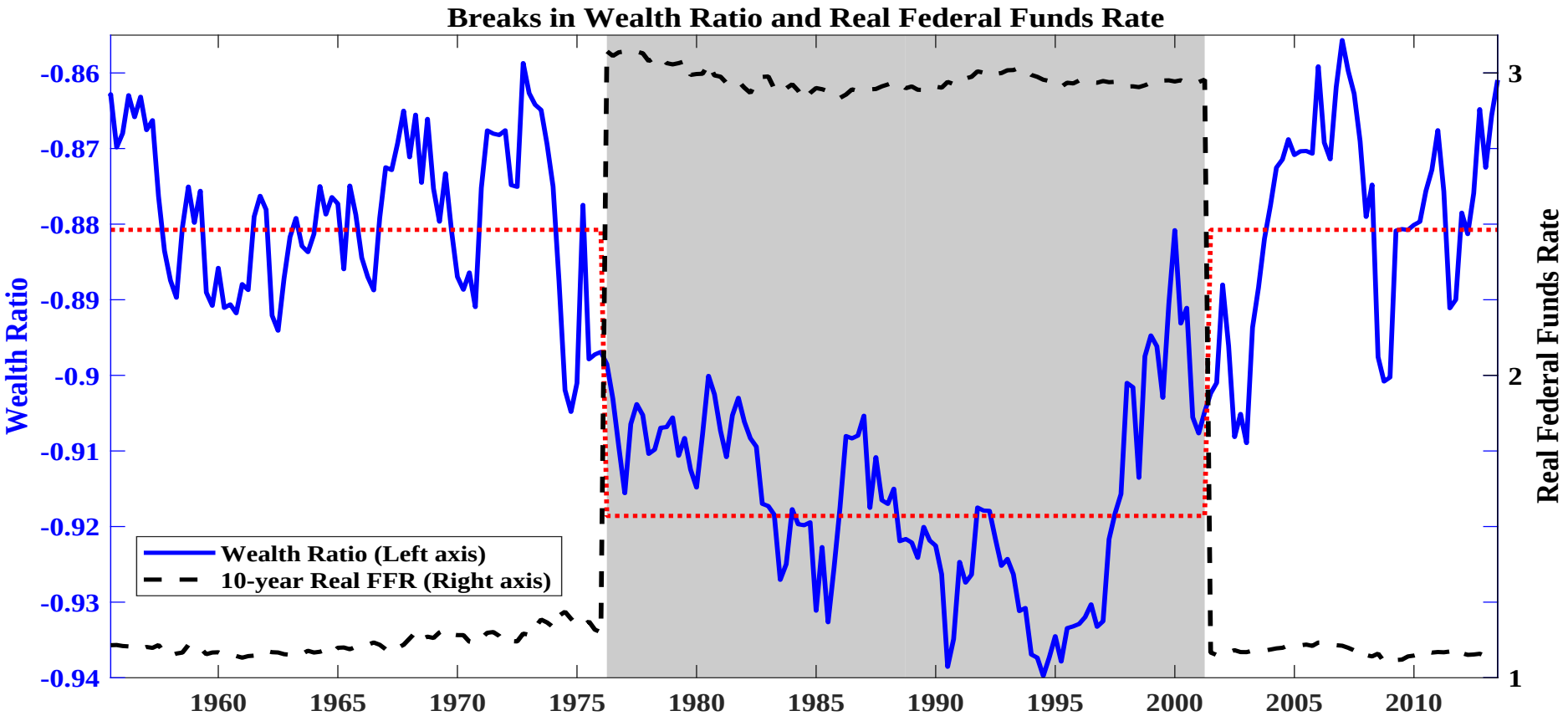



**Figure 4:** Wealth Ratio and federal funds rate. The wealth ratio (solid blue line, left axis) is plotted together with the conditional expectation of the ten-year-ahead real federal funds rate from the baseline MS-VAR (black dashed line, right axis). The wealth ratio is defined as the log inverse of $cay^{MS}$ without removing the Markov-switching constant. The red dashed line represents the log inverse of the regime-probability weighted average of the constants $\alpha_1$ and $\alpha_2$. The sample is quarterly and spans the period 1955:Q4 to 2013:Q3.

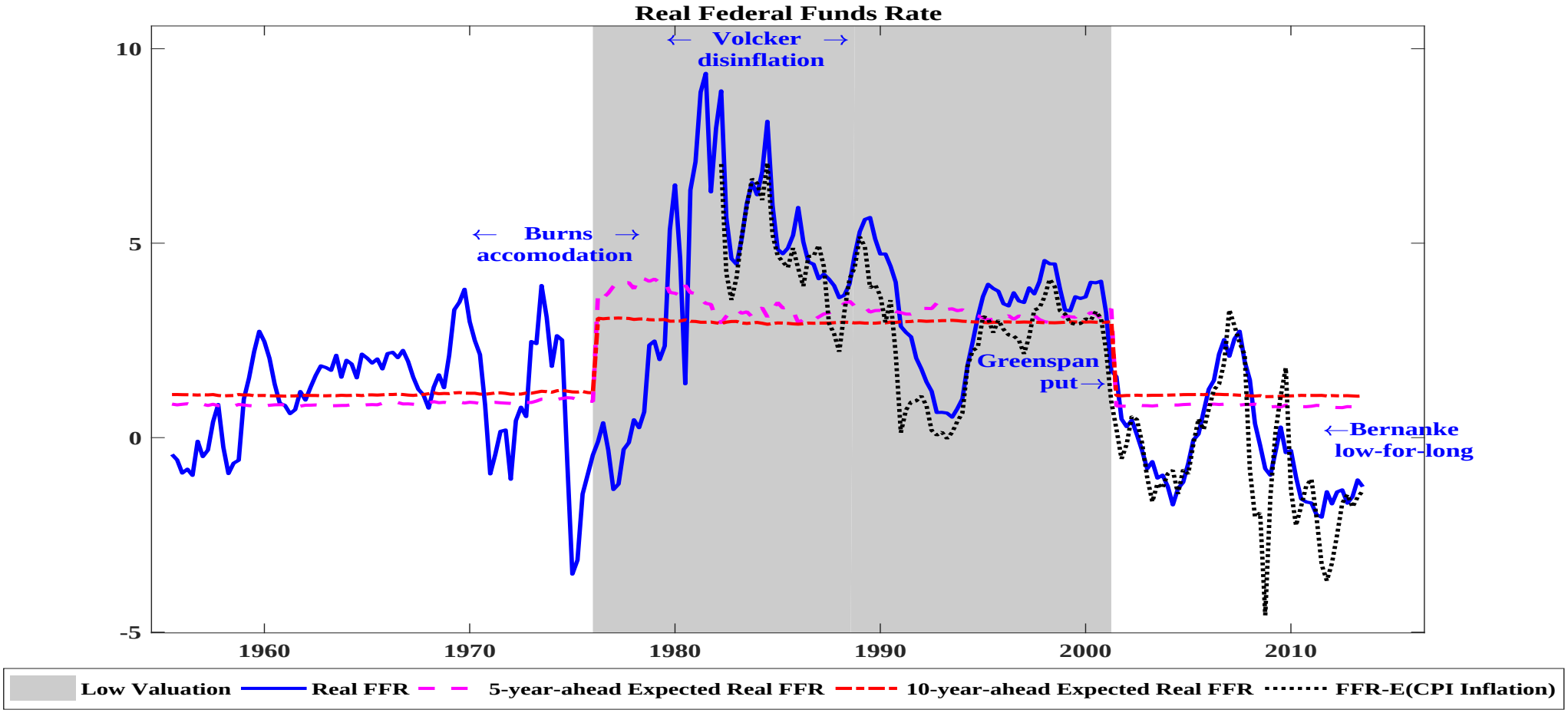



**Figure 5: Real Federal Funds Rate.** The figure reports the evolution of the Real FFR over time together with the 5-year-ahead and 10-year-ahead expectations as implied by the baseline MS-VAR. Expectations are computed taking into account the possibility of regime changes. The sample spans 1955:Q3 to 2013:Q3.

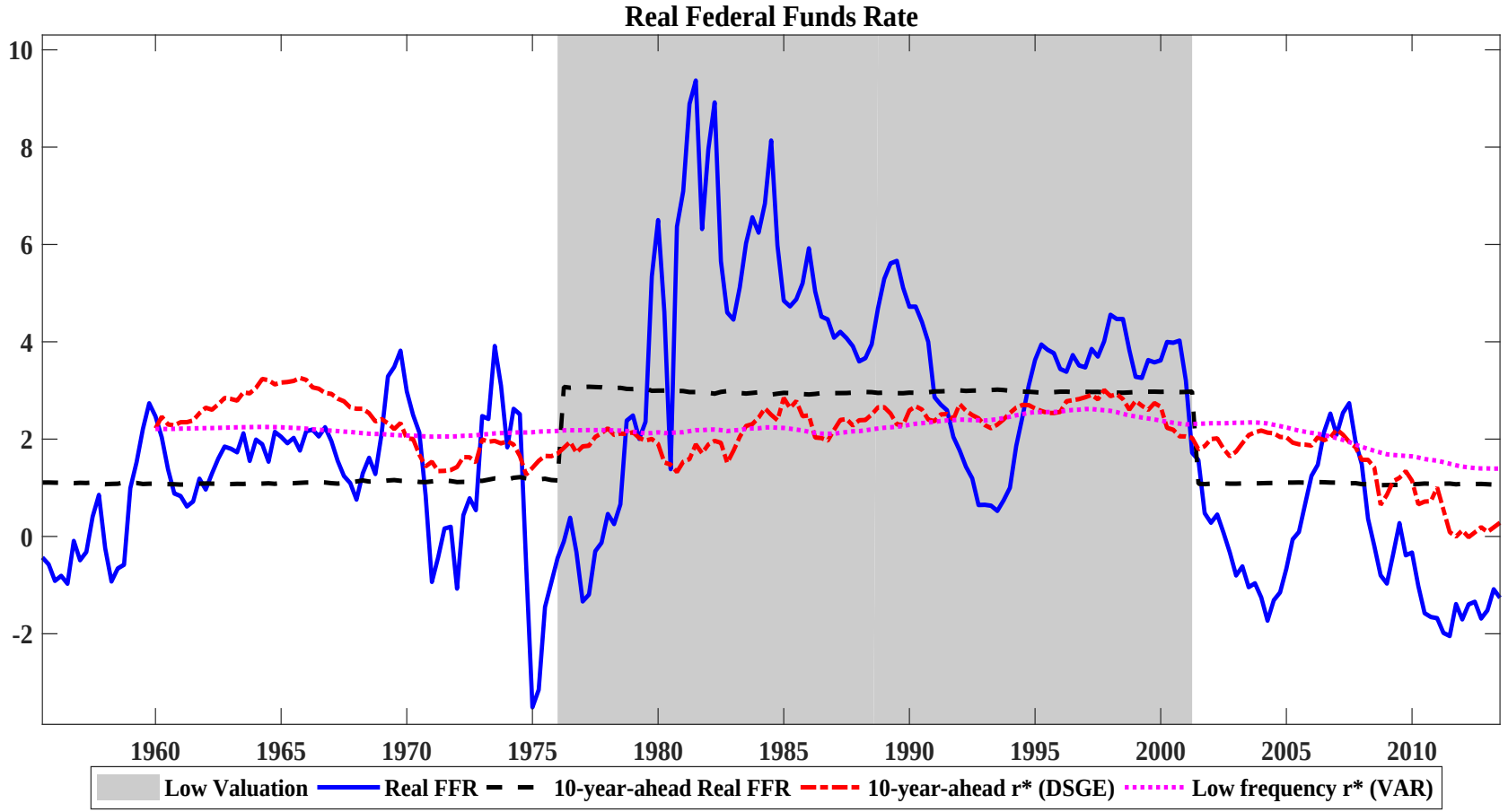



**Figure 6: Real Federal Funds Rate versus r*.** The figure reports the evolution of the Real FFR over time together with the expected 10-year-ahead from the baseline MS-VAR, and the 10-year-ahead forward $r^*$ from the DSGE model and VAR estimate of the low frequency component of $r^*$, both from Del Negro, Giannone, Giannoni, and Tambalotti (2017). The sample spans 1955:Q3 to 2013:Q3.

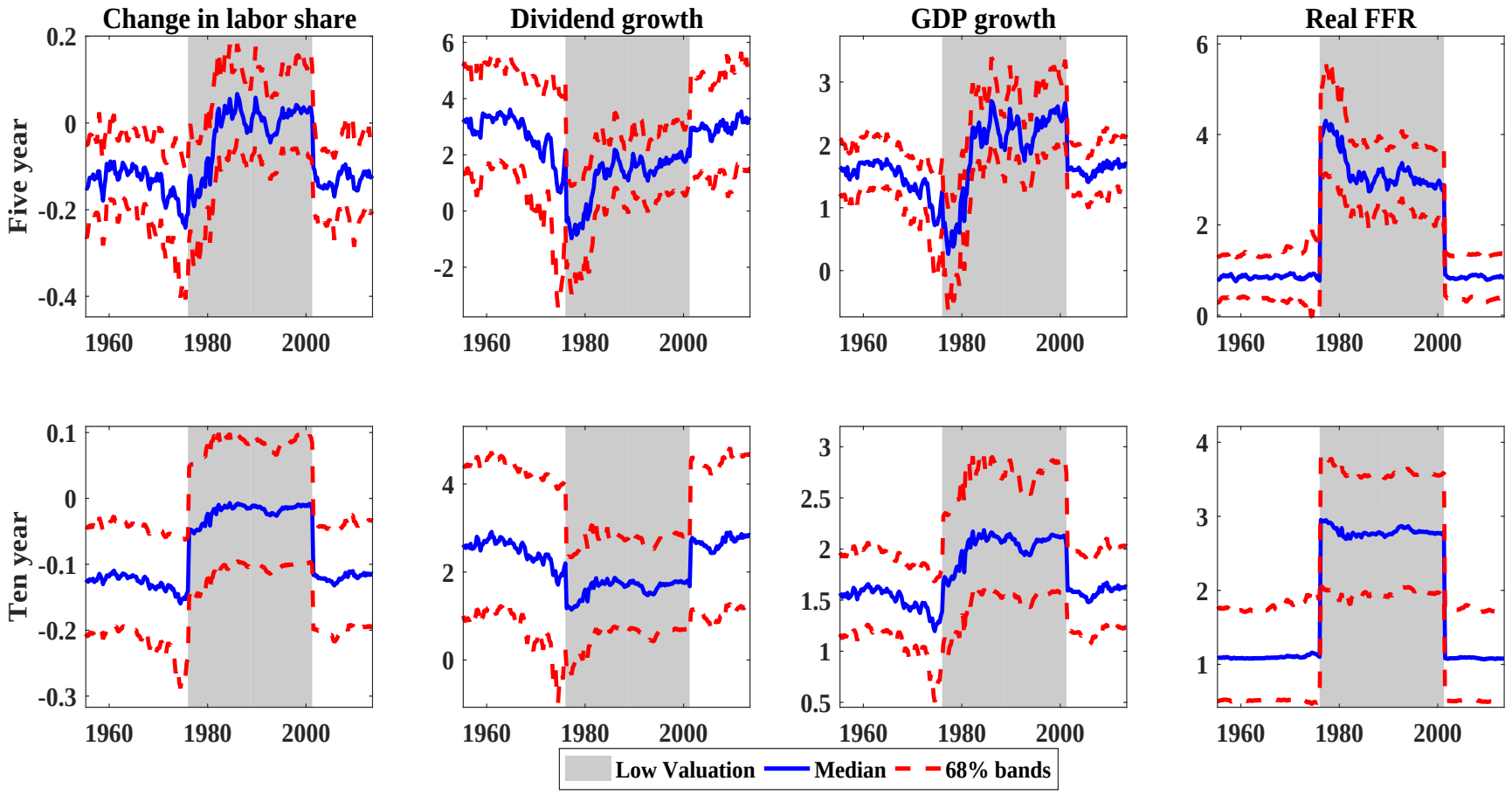



**Figure 7: Dividend growth and labor share.** The figure reports the conditional expectations for dividend growth, change in the labor share, GDP growth, and the real FFR. The expectations are based on a MS-VAR taking into account the possibility of regime changes. The sample spans 1955:Q3-2013:Q3.

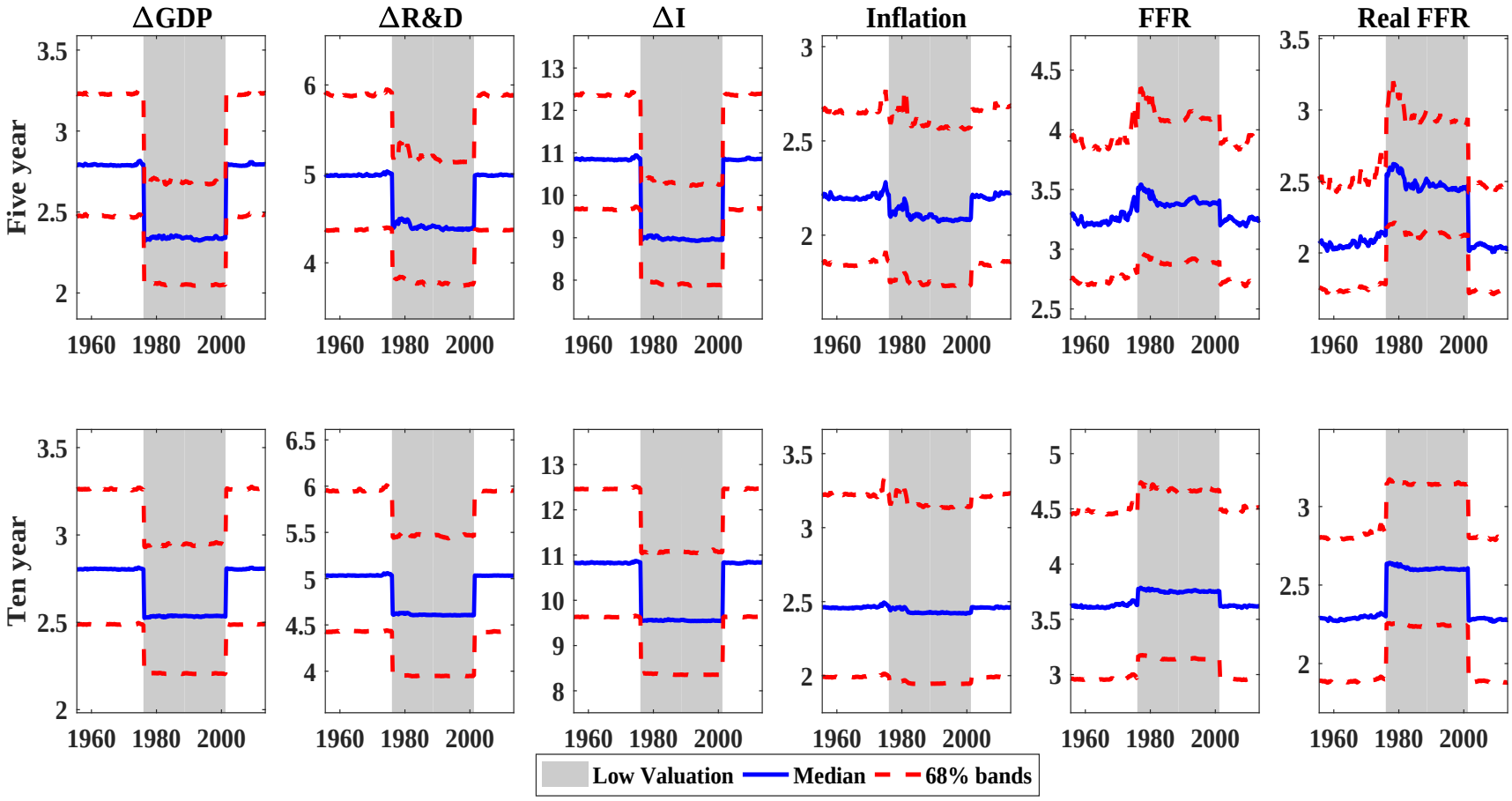



**Figure 8: Uncertainty based on MS-VAR.** The figure reports the conditional standard deviations at different horizons based on the MS-VAR taking into account the possibility of regime changes. The sample spans 1955:Q3-2013:Q3.

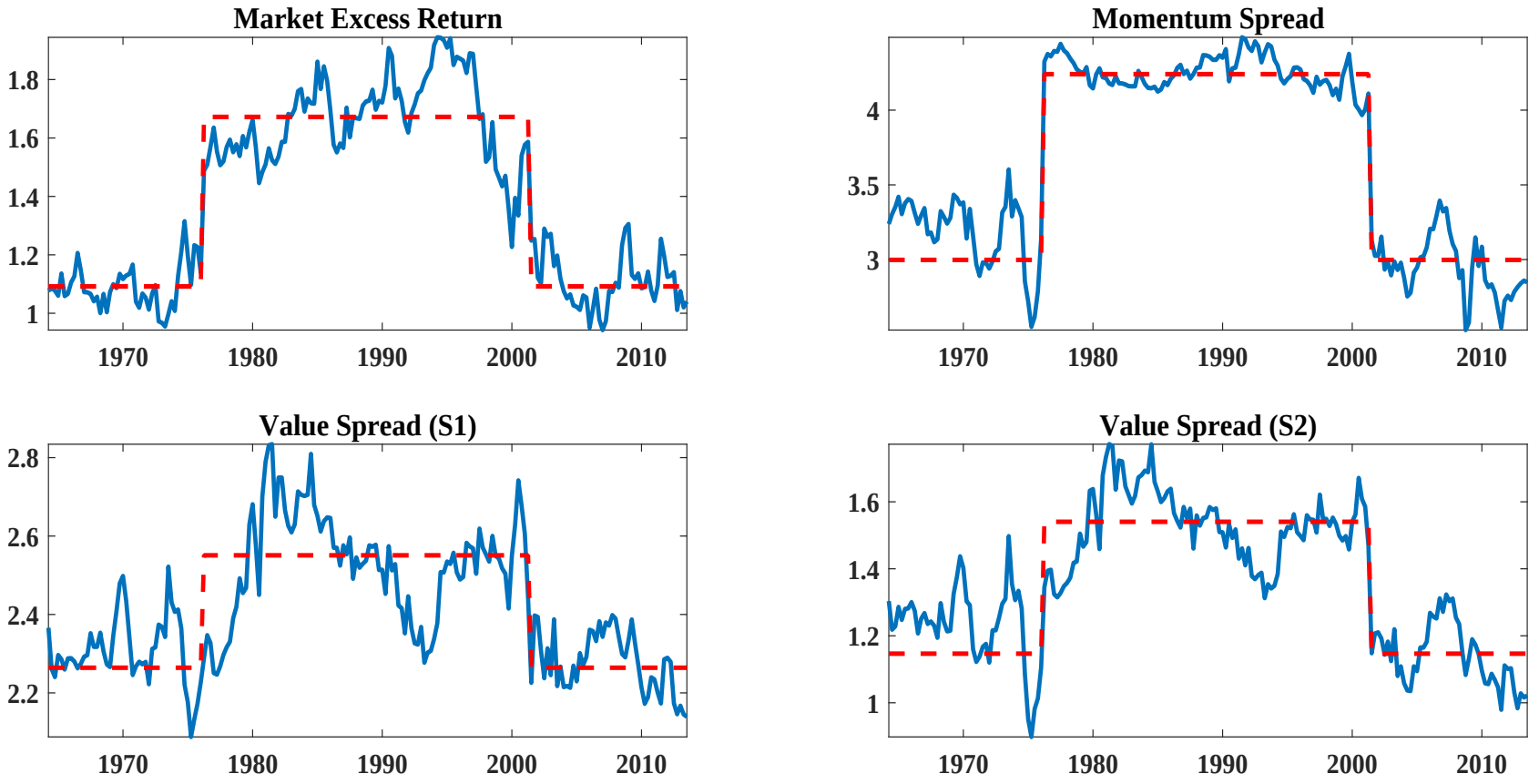



**Figure 9: Evolution of Risk Premia.** The figure reports the evolution of the PDV of risk premia for the stock market and three different spread portfolios. The blue solid line reports the evolution of the risk premia over time, while the red dashed line corresponds to the conditional steady state 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.

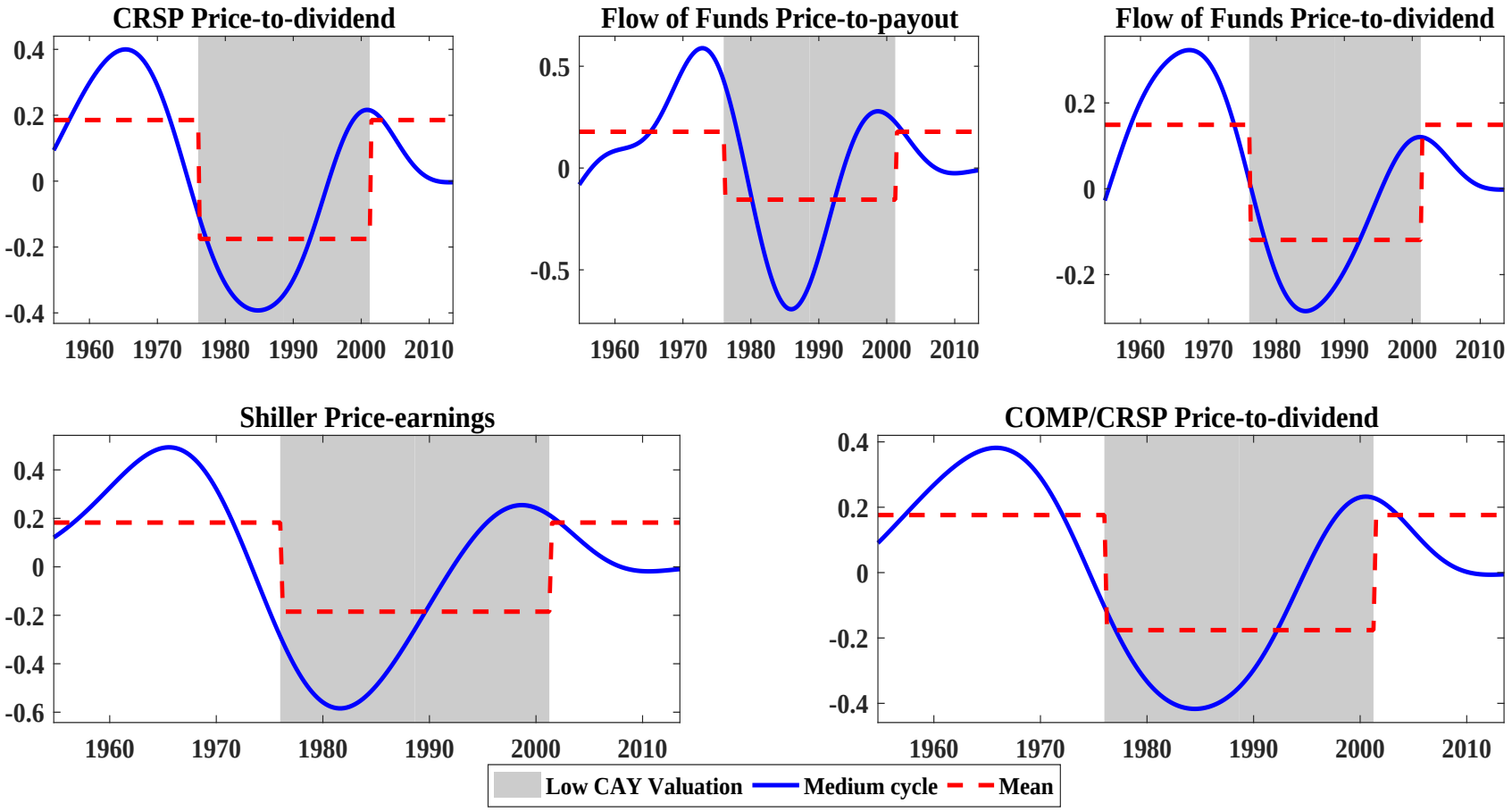



**Figure 10:** This figure presents the evolution of alternative valuation ratios at medium-term frequencies corresponding to cycles between 10 and 50 years. The sample spans 1955:Q3-2013:Q3.

# Appendix for Online Publication

## 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/.

CRSP PRICE-DIVIDEND RATIO

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

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

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

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

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 $\ln\left(P_t^{Ax}/D_t^A\right)$. Note that this value for dividend growth is only used to compute the CRSP price-dividend ratio on this hypothetical portfolio. When we investigate the behavior of stock market dividend growth in the MS-VAR, we use actual dividend data from all firms on NYSE, NASDAQ, and AMEX. See the data description for MS-VARs below.

FLOW OF FUNDS EQUITY PAYOUT, DIVIDENDS, PRICE

Flow of Funds payout is measured as "Net dividends plus net repurchases" and is computed using the Flow of Funds Table F.103 (nonfinancial corporate business sector) by subtracting Net Equity Issuance (FA103164103) from Net Dividends (FA106121075). We define net repurchases

to be repurchases net of share issuance, so net repurchases is the negative of net equity issuance. Net dividends consists of payments in cash or other assets, excluding the corporation's own stock, made by corporations located in the United States and abroad to stockholders who are U.S. residents. The payments are netted against dividends received by U.S. corporations, thereby providing a measure of the dividends paid by U.S. corporations to other sectors. The price used for FOF price-dividend and price-payout ratios is "Equity," the flow of funds measure of equities (LM103164103).

PRICE DEFLATOR FOR CONSUMPTION AND ASSET WEALTH

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.

DATA FOR MS-VARs

In the baseline MS-VAR, we use five observables: real R&D per capita growth, real investment per capita growth, real GDP per capita growth, annualized quarterly inflation, the federal funds rate. Our data sources are the NIPA tables constructed by the Bureau of Economic Analysis and the St. Louis Fed. Real GDP per capita is obtained by dividing nominal GDP (NIPA 1.1.5, line 1) by the GDP deflator (NIPA 1.1.4, line 1) and population. Consumption is defined as the sum of personal consumption expenditures on non-durable goods (NIPA 1.1.5, line 5) and services (NIPA 1.1.5, line 6). The series for nominal investment in physical capital is the sum of gross private domestic investment (NIPA 1.1.5, line 7) and personal consumption expenditure in durables (NIPA 1.1.5, line 4) *minus* intellectual property products (NIPA 1.1.5, line 12). Both series are then divided by the GDP deflator and population. Nominal R&D investment coincides with the series of intellectual property products (NIPA 1.1.5, line 12). The series is then divided by GDP deflator and population. Inflation is measured as the annual log-difference in the GDP deflator. The effective FFR is downloaded from the St. Louis Fed website, while all the other series are extracted from the NIPA tables constructed by the Bureau of Economic Analysis. The sample spans 1954:Q3 to 2013:Q3.

In the secondary MS-VAR that includes dividend growth and the change in the labor share, we compute dividends for all firms on NYSE, NASDAQ and AMEX using a merge of the Compustat Annual Industrial Database and the CRSP Monthly Stock Database, following the approach presented in Larrain and Yogo (2008). The labor share is defined as wages and salaries + employer contributions for employee pensions and insurance - employee contributions for social insurance. The quarterly data are in current dollars. Our source for these variables is the Bureau of Economic Analysis. Dividend growth is defined as the year-to-year change in log-dividends. The change in the labor share is computed by taking the year-to-year difference

in the annual mean of the labor share. These two variables are combined in the VAR with real GDP per capita growth, annualized quarterly inflation, and the effective federal funds rate.

## Computing $cay^{MS}$

Let $\boldsymbol{z}_t$ be a $3\times1$ vector of data on $c_t$, $a_t$, and $y_t$, and $k$ leads and $k$ lags of $\Delta a_t$ and $\Delta y_t$ 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 dynamic least squares regression (DLS–Stock and Watson (1993)) regression to allow for shifts in the intercept $\alpha_{\xi_t^\alpha}$:

$$c_t = \alpha_{\xi_t^\alpha} + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i}\Delta a_{t+i} + \sum_{i=-k}^{k} b_{y,i}\Delta y_{t+i} + \sigma\varepsilon_t \tag{A1}$$

where $\varepsilon_t \sim N(0,1)$.[23] 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 $\alpha_1$ and $\alpha_2$, the standard deviation of the residual $\sigma$, and the transition probabilities contained in the matrix $\boldsymbol{H}^\alpha$. Collect these parameters into a vector $\boldsymbol{\theta} = \left(\beta, \alpha_{\xi_t^\alpha}, \sigma, H^\alpha\right)'$.

Our estimate of $cay_t^{MS}$ is based on the posterior mode of the parameter vector $\boldsymbol{\theta}$ and the corresponding regime probabilities. Collect the conditional probabilities $\pi_{t|t}^i = p(\xi_t^\alpha = i|Y^t;\boldsymbol{\theta})$ for $i = 1, .., m$ into an $m \times 1$ vector $\pi_{t|t} = p(\xi_t^\alpha|Y^t;\boldsymbol{\theta})$. 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 given by the Hamilton filter:

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

where $\eta_t$ 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 $\xi_0^\alpha$. We assume that the two regimes have equal probabilities: $p(\xi_0^\alpha = 1) = .5 = p(\xi_0^\alpha = 2)$.

[23] The DLS regression controls for leads and lags of the right-hand-side variables to adjust for the inefficiencies attributable to regressor endogeneity that arise in finite samples.

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 (A1) 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 $\boldsymbol{H}^{\alpha}$.

**Regime probabilities**

This section presents additional details about the regime probabilities. Regime probabilities are based on the DLS regression with Markov-switching constant presented in (A1). Thus what matters to establish which regime is in place is the relative size of the residuals from (A1) under the two regimes, not the residual of the cointegrating relation. In a finite sample, the residuals from the MS-DLS regression and the cointegrating residuals would coincide only if leads and lags were not included.

Figure A.1 zooms in on the filtered and smoothed regime probabilities around the dates of the regime changes. This figure confirms that there is little uncertainty about the timing of the regime changes. To understand why, it is useful to analyze the behavior of the residuals of the Markov-switching DLS regression across the two regimes. These are presented in the left panel of Figure A.2. The right panel of the same figure presents the squared values of the residuals. As explained in the previous subsection, the regime probabilities largely depend on the squared residuals across the two regimes. Of course, the transition matrix matters too, but in this case its role is marginal. Note that there is a clear jump in the relative size of the squared residuals around the dates that we have identified for the regime changes. This is important, for example, to understand why the peak of asset valuation at the end of the 1990s is included in the low valuation regime: When looking of the residuals of the DLS regression it is clear that the end of

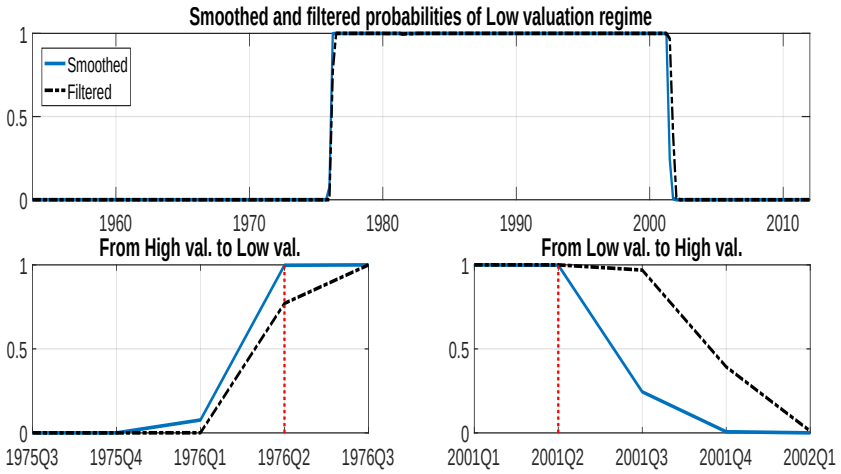



**Figure A.1:** Filtered and smoothed regime probabilities around switch dates.

the 1990s are part of the low valuation regime. This can be rationalized in light of the strong macroeconomic fundamentals of these years and the fact that the run-up in asset valuations was in fact quite short lasting. Our regime changes capture changes in valuation that occur at lower frequencies. It is interesting to notice that it is exactly because of these properties that our model correctly capture that the end of the 1990s are followed by a sharp decline of asset valuations and then a steady increase once the economy enters the high valuation regime.

## Gibbs Sampling Algorithm

This appendix describes the Bayesian methods used to characterize uncertainty in the regression parameters. 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 $\beta_r$, $\alpha_{\xi_t^\alpha,r}$, $\sigma_r$, $H_r^\alpha$, 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** $\beta_{r+1}$: Given $\alpha_{\xi_t^\alpha,r}$, $\sigma_r$, and $\xi_r^{\alpha,T}$ we transform the data:

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

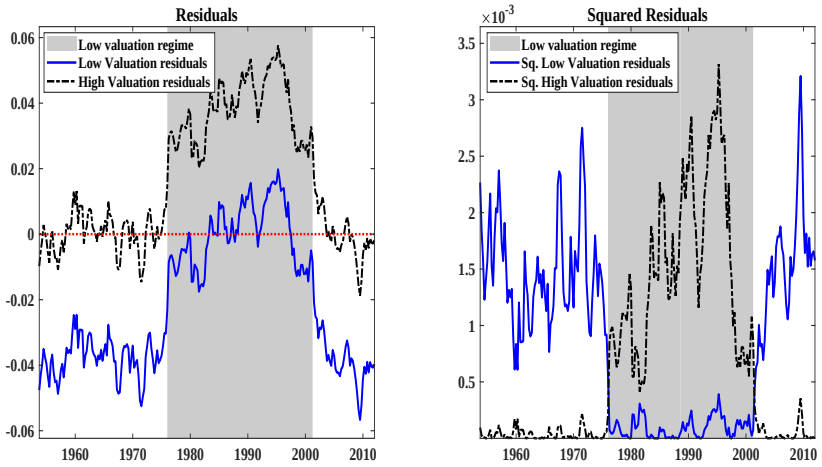



**Figure A.2:** The figure reports the behavior of the residuals (left panel) and the squared residuals (right panel) of the Markov-switching DLS regression across the two regimes.

The above is a regression with fixed coefficients $\beta$ 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 $\alpha_{\xi_t^\alpha,r}$, $\sigma_r$, 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 $\beta_{r+1}$, $\sigma_r$, 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 $\alpha_i$ from the conditional (on $\beta_{r+1}$, $\sigma_r$, 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** $\sigma_{r+1}$: Given $\beta_{r+1}$, $\alpha_{\xi_t^\alpha,r+1}$, 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 $\sigma$ has an inverse gamma distribution, $\sigma \sim IG\left(Q_0, v_0\right)$, we use Bayesian methods to draw $\sigma_{r+1}$ from the conditional (on $\beta_{r+1}$, $\alpha_{\xi_t^\alpha,r+1}$, 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 $\sigma$ is therefore $\left(E'E\right)/T$, which coincides with the standard maximum likelihood (MLE) estimate of $\sigma$, conditional on the other parameters.

4. **Sampling** $\xi_{r+1}^{\alpha,T}$: Given $\beta_{r+1}$, $\alpha_{\xi_t^\alpha,r+1}$, and $H_r^\alpha$ 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** $H_{r+1}^\alpha$: 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.[24] For each column of $H_{r+1}^\alpha$ 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 $i^\alpha$ to state $j^\alpha$ based on $\xi_{r+1}^{\alpha,T}$, while $a_{ii}^\alpha$ and $a_{ij}^\alpha$ 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.

[24] 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. 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 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.

**Convergence Checks**

The 90% credible sets are obtained making 2,000,000 draws from the posterior using the Gibbs sampling algorithm. One in every one thousand draws is retained. We check convergence using the methods suggested by Raftery and Lewis (1992) and Geweke (1992). The results are reported in Table A.1 and Table A.2, respectively. For Raftery and Lewis (1992) checks, we target 90% credible sets, with a 1% accuracy to be achieved with a 95% minimum probability. 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. 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 |
|---|---|---|---|---|---|---|---|---|
| $\alpha_1$ | 17413 | 9.541 | $\Delta a_{t+1}$ | 1799 | 0.986 | $\Delta y_{t-4}$ | 1850 | 1.014 |
| $\alpha_2$ | 16949 | 9.287 | $\Delta y_{t+1}$ | 1812 | 0.993 | $\Delta a_{t+4}$ | 1793 | 0.982 |
| $\beta_a$ | 1918 | 1.051 | $\Delta a_{t-2}$ | 1830 | 1.003 | $\Delta y_{t+4}$ | 1820 | 0.997 |
| $\beta_y$ | 1843 | 1.01 | $\Delta y_{t-2}$ | 1801 | 0.987 | $\Delta a_{t-5}$ | 1797 | 0.985 |
| $\sigma$ | 1797 | 0.985 | $\Delta a_{t+2}$ | 1886 | 1.033 | $\Delta y_{t-5}$ | 1850 | 1.014 |
| $H_{11}^{\alpha}$ | 1826 | 1.001 | $\Delta y_{t+2}$ | 1767 | 0.968 | $\Delta a_{t+5}$ | 1826 | 1.001 |
| $H_{22}^{\alpha}$ | 1820 | 0.997 | $\Delta a_{t-3}$ | 1858 | 1.018 | $\Delta y_{t+5}$ | 1850 | 1.014 |
| $\Delta a_t$ | 1823 | 0.999 | $\Delta y_{t-3}$ | 1808 | 0.991 | $\Delta a_{t-6}$ | 1850 | 1.014 |
| $\Delta y_t$ | 1850 | 1.014 | $\Delta a_{t+3}$ | 1847 | 1.012 | $\Delta y_{t-6}$ | 1826 | 1.001 |
| $\Delta a_{t-1}$ | 1839 | 1.008 | $\Delta y_{t+3}$ | 1820 | 0.997 | $\Delta a_{t+6}$ | 1839 | 1.008 |
| $\Delta y_{t-1}$ | 1866 | 1.022 | $\Delta a_{t-4}$ | 1830 | 1.003 | $\Delta y_{t+6}$ | 1866 | 1.022 |

**Table A.1:** 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.

| Variable | NSE | RNE | Variable | NSE | RNE | Variable | NSE | RNE |
|---|---|---|---|---|---|---|---|---|
| $\alpha_1$ | 0.000131 | 1 | $\Delta a_{t+1}$ | 0.000263 | 1 | $\Delta y_{t-4}$ | 0.000526 | 1 |
| $\alpha_2$ | 0.000131 | 1 | $\Delta y_{t+1}$ | 0.00053 | 1 | $\Delta a_{t+4}$ | 0.000256 | 1 |
| $\beta_a$ | 0.000074 | 1 | $\Delta a_{t-2}$ | 0.000261 | 1 | $\Delta y_{t+4}$ | 0.000521 | 1 |
| $\beta_y$ | 0.000085 | 1 | $\Delta y_{t-2}$ | 0.000572 | 1 | $\Delta a_{t-5}$ | 0.000264 | 1 |
| $\sigma$ | 0 | 1 | $\Delta a_{t+2}$ | 0.000258 | 1 | $\Delta y_{t-5}$ | 0.000524 | 1 |
| $H_{11}^{\alpha}$ | 0.000069 | 1 | $\Delta y_{t+2}$ | 0.000547 | 1 | $\Delta a_{t+5}$ | 0.000252 | 1 |
| $H_{22}^{\alpha}$ | 0.000053 | 1 | $\Delta a_{t-3}$ | 0.000278 | 1 | $\Delta y_{t+5}$ | 0.000534 | 1 |
| $\Delta a_t$ | 0.000263 | 1 | $\Delta y_{t-3}$ | 0.000632 | 1 | $\Delta a_{t-6}$ | 0.000275 | 1 |
| $\Delta y_t$ | 0.000529 | 1 | $\Delta a_{t+3}$ | 0.000255 | 1 | $\Delta y_{t-6}$ | 0.000518 | 1 |
| $\Delta a_{t-1}$ | 0.000252 | 1 | $\Delta y_{t+3}$ | 0.000537 | 1 | $\Delta a_{t+6}$ | 0.000238 | 1 |
| $\Delta y_{t-1}$ | 0.000521 | 1 | $\Delta a_{t-4}$ | 0.000259 | 1 | $\Delta y_{t+6}$ | 0.000525 | 1 |

**Table A.2:** 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.

## Model Comparison

This section reports the results of tests that compare competing models for $cay_t^{MS}$. We analyze two alternative models and compare them to our two benchmark models for $cay$: $cay^{MS}$ in

which only the constant is allowed to switch over time, and $cay^{FC}$ in which there is no Markov-switching in any of the parameters. In the first alternative, we allow for Markov-switching in heteroskedasticity as well as Markov-switching in the constant. In the second alternative, we only allow for Markov-switching in only heteroskedasticity. We use the Bayesian information criterion ($BIC$) to compare these different models. This is computed as:

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

where $maxli$ is the maximized *log*-likelihood, $k$ is the number of parameters, and $T$ the sample size. A model with a smaller $BIC$ criterion is preferred to one with a larger criterion. The $BIC$ criterion rewards models with higher likelihoods but also penalizes models that have more parameters.

Table A.3 reports the estimates for the key *cay* parameters and the $BIC$ for each model. The $BIC$ is lowest by the model is the one that allows for switching in both heteroskedasticity and the constant (MS $\alpha$ and MS $\sigma$). But our benchmark model with switches only in the constant (MS $\alpha$ only) is preferred to the model that with switches only in heteroskedasticity (MS $\sigma$ only), and to the fixed coefficient model (FC). These results support the hypothesis of shifts in the constant. And although the model with switching in both the constant and the volatility is preferred, the estimates for the cointegrating vector and the timing of regimes are essentially 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 | $\alpha_1$ | $\alpha_2$ | $\beta_a$ | $\beta_y$ | $\sigma_1$ | $\sigma_2$ | $BIC$ |
|---|---|---|---|---|---|---|---|
| MS $\alpha$ and MS $\sigma$ | 0.9186 | 0.8810 | 0.2599 | 0.6162 | 0.0016 | 0.0105 | −1472.0 |
| MS $\alpha$ only | 0.9186 | 0.8808 | 0.2606 | 0.6156 | 0.0080 | | −1390.3 |
| MS $\sigma$ only | 0.8056 | | 0.1275 | 0.7845 | 0.0029 | 0.0204 | −1230.2 |
| FC | 0.8706 | | 0.1246 | 0.7815 | 0.0158 | | −1202.2 |

**Table A.3:** 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 $\alpha$ and MS $\sigma$: The model allows for changes in the constant and heteroskedasticity. MS $\alpha$ only: Benchmark model with only changes in the constant. MS $\sigma$ only: The model allows for heteroskedasticity, but not changes in the constant. FC: Standard fixed coefficient regression.

## Persistence of $\epsilon^{MS}$ versus $\epsilon^{FC}$

The column of Table A.4 reports the first-order autoregressive coefficient estimate for the two versions of *cay*. The estimated autocorrelation coefficient for $\epsilon_t^{FC}$ is 0.94. The estimated first-

order autocorrelation coefficient for $\epsilon_t^{MS}$ is 0.81.

**Cointegration Tests**

| | | *Dickey–Fuller t-statistic* | | | | *Critical values* | |
|---|---|---|---|---|---|---|---|
| | *Persistence* $\epsilon$ | $Lag = 1$ | $Lag = 2$ | $Lag = 3$ | $Lag = 4$ | 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 A.4:** 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 the Markov-switching version 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.

Several other tests are employed to assess the degree of persistence in $\epsilon_t^{MS}$ as compared to $\epsilon_t^{FC}$. First, we apply an augmented Dickey-Fuller $t$ test to the estimated cointegrating residuals. The test statistics and corresponding critical values are reported in Table A.4. According to this test, the null hypothesis of no cointegration is rejected for $\epsilon_t^{MS}$ in every case, whereas the opposite is true for $\epsilon_t^{FC}$.

Second, we examine low frequency averages or "cosine transformations" of $\epsilon$ to gauge its persistence following Muller and Watson (2008) and Watson (2013). The cosine transformation of $\epsilon_t^{MS}$ displays a pattern much more consistent with an $I(0)$ series than that of $\epsilon_t^{FC}$. Figure A.3 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 $\epsilon$

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

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 A.3 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 A.3 shows that the cosine transformation of $\epsilon_t^{MS}$ displays a pattern much more consistent with an $I(0)$ series than that of $\epsilon_t^{FC}$, which shows a clear spike at $j = 3$, corresponding to a period of roughly 41 years.

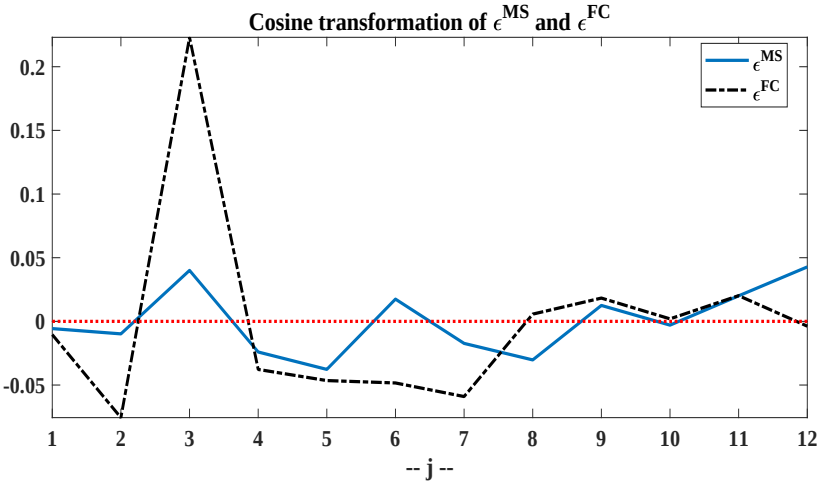



**Figure A.3:** Low frequency averages of $\epsilon^{MS}$ and $\epsilon^{FC}$. The figure plots the set of averages $\{f_j\}_{j=1}^{k}$, which capture the variability in *cay* for periods greater than $2T/k$, where $T$ is the sample size. Thus, with T = 247 quarters, the $k = 12$ points plotted summarize the variability in $\epsilon^{MS}$ and $\epsilon^{FC}$ 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.

Third, we estimate fractionally integrated models for $\epsilon_t^{MS}$ and $\epsilon_t^{FC}$ in which $(1-L)^d \epsilon_t^{(m)} = u_t$, where $L$ is the lag operator, $u_t$ is an $I(0)$ process and $m = MS, FC$. 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:

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

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 $\epsilon$, 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:

$$p_{X^S}(x^s) = \frac{1}{2}\Gamma(k/2)\, \pi^{-k/2} |\Sigma_X|^{-1/2} \left(x^{s\prime} \Sigma_X^{-1} x^s\right)^{-q/2} \tag{A2}$$

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

We estimate a fractionally integrated model for $cay_t$: $(1-L)^d \epsilon_t^{(m)} = u_t$, 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:

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

where $m = MS, FC$. Note that when $d = 1$, the fractional integrated model implies that $cay_t$ has a unit root, $\epsilon_t = \epsilon_{t-1} + u_t$, while for $d = 0$, $\epsilon_t = u_t$, i.e. $\epsilon_t$ 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 $\Psi$ 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 (A2) 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. If $\epsilon_t^{(m)}$ is $I(0)$, then $d = 0$. If it has a unit root, then $d = 1$, and non-integer values of $d > 0$ are fractionally integrated series that are more persistent than $I(0)$ but less persistent than $I(1)$. Figure A.4 shows the estimated log likelihoods for $(1-L)^d \epsilon_t^{MS}$ and $(1-L)^d \epsilon_t^{FC}$ as a function of $d$. For $\epsilon_t^{MS}$, the likelihood peaks at $d = 0$, while for $\epsilon_t^{FC}$, the likelihood rises with $d > 0$ and peaks near $d = 1.2$.

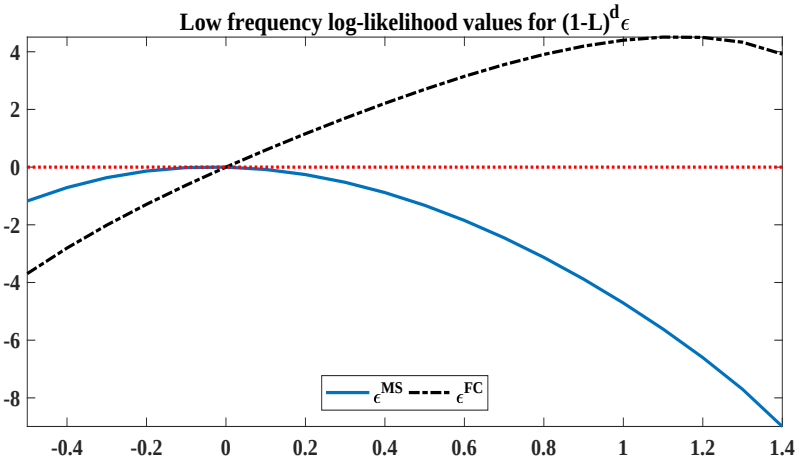



**Figure A.4:** Low frequency log likelihood values for $(1-L)^d \epsilon^{MS}$ and $(1-L)^d \epsilon^{FC}$. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

## MS-VAR Estimation

In this appendix we provide details on the estimation of the MS-VAR 5. 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

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

where $\eta_t$ 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

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

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 $\varepsilon$ is $\varepsilon_t$, and where $\Sigma_{\xi_t} = V_{\xi_t} V'_{\xi_t}$. If we specify a Normal-Wishart prior for $A_{\xi_t}$ and $V_{\xi_t}$:

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

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 $A_{\xi_t}$ 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.[25] 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 $\boldsymbol{H}^A$.

## Conditional 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:

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t} \varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{A3}$$

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:

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

[25] 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).

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:

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

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:

**Proposition 1** *Consider a Markov-switching model whose law of motion can be described by (A3) 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 $\mu_{t+s|t} \equiv \mathbb{E}_t(Z_{t+s})$, 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

$$\mu_{t+s|t} = \mathbb{E}_t(Z_{t+s}) = w q_{t+s|t}, \tag{A4}$$

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

$$q_{t+s|t} = C\pi_{t+s|t} + \Omega q_{t+s-1|t} \tag{A5}$$

$$\pi_{t+s|t} = H\pi_{t+s-1|t} \tag{A6}$$

with $\pi_{t+s|t} = \left[\pi_{t+s|t}^1, ..., \pi_{t+s|t}^m\right]'$, $\Omega = bdiag(A_1, ..., A_m)(H \otimes I_n)$, and $\underset{mn\times m}{C} = bdiag(c_1, ..., c_m)$, 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 $\varphi(X)$ 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: $\varphi(X_1 X_2 X_3) = (X_3' \otimes X_1)\varphi(X_2)$.

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

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

where the matrix $W = [I_{n^2}, ..., I_{n^2}]$ 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 (A3) and define* $Q_t^i = \varphi\left[\mathbb{E}_0\left(Z_t Z_t' 1_{\xi_t=i}\right)\right]$, $q_t^i = \mathbb{E}_0\left[Z_t 1_{\xi_t=i}\right]$, *and* $\pi_t^i = P_0\left(\xi_t = i\right)$, *for* $i = 1...m$. *Then* $Q_t^j = \left[\widehat{cc}_j + \widehat{VV}_j \varphi\left[I_k\right]\right]\pi_t^j + \sum_{i=1}^m \left[\widehat{AA}_j Q_{t-1}^i + \widehat{DAC}_j q_{t-1}^i\right] h_{ji}$, *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:

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

where $Q_{t+s|t}^i = \varphi\left(\mathbb{E}_t\left(Z_{t+s} Z_{t+s}' 1_{\xi_{t+s}=i}\right)\right) = \varphi\left(\mathbb{E}_t\left(Z_{t+s} Z_{t+s}' | \xi_{t+s} = i\right)\right)\pi_{t+s|t}^i$, 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:

$$\begin{aligned} Q_{t+s|t} &= \Xi Q_{t+s-1|t} + \widehat{DAC} q_{t+s-1|t} + \widehat{Vc}\pi_{t+s|t} && \text{(A7)} \\ q_{t+s|t} &= C\pi_{t+s|t} + \Omega q_{t+s-1|t},\ \pi_{t+s|t} = H\pi_{t+s-1|t}. && \text{(A8)} \end{aligned}$$

where

$$\begin{aligned} \Xi &= bdiag(\widehat{AA}_1, ..., \widehat{AA}_m)(H \otimes I_{n^2}),\ \widehat{Vc} = \left[\widehat{VV} + \widehat{cc}\right],\ \widehat{cc} = bdiag(\widehat{cc}_1, ..., \widehat{cc}_m), \\ \widehat{VV} &= bdiag(\widehat{VV}_1 \varphi\left[I_k\right], ..., \widehat{VV}_m \varphi\left[I_k\right]),\ \widehat{DAC} = bdiag(\widehat{DAC}_1, ..., \widehat{DAC}_m)(H \otimes I_n). \end{aligned}$$

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 $t$:

$$\varphi\left[\mathbb{V}_t\left(Z_{t+s}\right)\right] = M_{t+s|t} - \varphi\left[\mu_{t+s|t}\mu_{t+s|t}'\right], \tag{A9}$$

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

To report estimates of (A4) and (A9) 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 (A5)-(A6) and (A7)-(A8) in a compact form as

$$\widetilde{Q}_{t+s|t} = \widetilde{\Xi}^s \widetilde{Q}_{t|t} \text{ where } \widetilde{\Xi} = \left[\begin{array}{c|cc} \Xi & \widehat{DAC} & \widehat{Vc}H \\ \hline & \Omega & CH \\ & & H \end{array}\right], \tag{A10}$$

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 (A4) and (A9) using (A10). 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, $\pi_{t|t}^i \equiv P\left(\xi_t = i | \mathbb{I}_t\right)$ 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 $Z_t \in \mathbb{I}_t$, $q'_{t|t} = \left[q^{1'}_{t|t}, q^{2'}_{t|t}\right]'$ 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 $Q^i_{t|t} \equiv \varphi\left(\mathbb{E}_t\left(Z_t Z'_t|\xi_t = i\right)\right)\pi^i_{t|t}$ will be $\left[\varphi\left(Z_t Z'_t \cdot 1\right)', \varphi\left(Z_t Z'_t \cdot 0\right)'\right]'$ or $\left[\varphi\left(Z_t Z'_t \cdot 0\right)', \varphi\left(Z_t Z'_t \cdot 1\right)'\right]'$.

## Mean Square Stability

When estimating the MS-VAR we require the model to be mean square stable. Mean square stability is defined as follows:

**Definition 1** *An n-dimensional process $Z_t$ is mean square stable if and only if there exists an n-vector $\overline{\mu}$ and an $n^2$-vector $\overline{M}$ such that:*

*1)* $\lim_{t\to\infty} \mathbb{E}_0\left[Z_t\right] = \overline{\mu}$

*2)* $\lim_{t\to\infty} \mathbb{E}_0\left[Z_t Z'_t\right] = \overline{M}$

*for any initial $Z_0$ and $\xi_0$.*

Mean-square-stability requires that first and second moments converge as the time horizon goes to $\infty$. Under the assumptions that the Markov-switching process $\xi_t$ is ergodic and that the innovation process $\varepsilon_t$ is asymptotically covariance stationary, Costa, Fragoso, and Marques (2004) show that a multivariate Markov-switching model as the one described by (5)-(**??**) is mean-square stable if and only if it is asymptotically covariance stationary. Both conditions hold for the models studied in this paper and are usually verified in economic models.

Costa, Fragoso, and Marques (2004) show that in order to establish MSS of a process such as the one described by (5)-(**??**), it is enough to check MSS stability of the correspondent homogeneous process: $Z_t = A_{\xi_t} Z_{t-1}$. In this case, formulas (**??**) and (A10) simplify substantially: $q_t = \Omega q_{t-1}$ and $Q_t = \Xi Q_{t-1}$. Let $r_\sigma\left(X\right)$ be the operator that given a square matrix $X$ computes its largest eigenvalue. We then have:

**Proposition 3** *A Markov-switching process whose law of motion can be described by (5)-(**??**) is mean square stable if and only if $r_\sigma\left(\Xi\right) < 1$.*

Mean square stability allows us to compute finite measures of uncertainty as the time horizon goes to infinity. Mean square stability also implies that shocks do not have permanent effects on the variables included in the MSVAR.

## Long Term Response of the FFR

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

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

where $Z_t$ is an $n \times 1$ vector of variables, $c_{\xi_t}$ is an $n \times 1$ vector of constants, $A_{l,\xi_t}$ 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 shocks $\varepsilon_t$. The process $\xi_t$ 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 $cay^{MS}$.

In order to obtain the structural representation of the MS-VAR we are going to decompose the covariance matrix $V_{\xi_t} V'_{\xi_t}$ as $B^{-1}_{0,\xi_t} W_{\xi_t} W'_{\xi_t} B^{-1}_{0,\xi_t}$ for each regime. The matrix $B^{-1}_{0,\xi_t}$ captures the contemporaneous relations among the variables of interest, while the matrix $W_{\xi_t}$ is the covariance matrix for the structural disturbances. We use a Cholesky decomposition to identify the contemporaneous response of the FFR to the other variables. This identifying assumption implies that the Federal Reserve can react contemporaneously to all variables included in our estimates, while the rest of the variables react with a lag. We do not attempt to identify other shocks, so the triangular structure imposed on the remaining variables should be interpreted as a normalization as opposed to an identifying restriction.

We then have:

$$Z_t = c_{\xi_t} + A_{1,\xi_t} Z_{t-1} + A_{2,\xi_t} Z_{t-2} + B^{-1}_{0,\xi_t} W_{\xi_t} \varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{A11}$$

$$B_{0,\xi_t} Z_t = B_{0,\xi_t} c_{\xi_t} + B_{0,\xi_t} A_{1,\xi_t} Z_{t-1} + B_{0,\xi_t} A_{2,\xi_t} Z_{t-2} + W_{\xi_t} \varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{A12}$$

This can rewritten as:

$$B_{0,\xi_t} Z_t = B_{0,\xi_t} c_{\xi_t} + B_{1,\xi_t} Z_{t-1} + B_{2,\xi_t} Z_{t-2} + \omega_t,\ \omega_t \sim N\left(0, W_{\xi_t} W'_{\xi_t}\right)$$

Suppose we are interested in the long term response of the FFR to inflation. We then extract the relevant equation for the FFR and focus on the effects coming just from inflation:

$$\begin{aligned} FFR_t &= -B_{0,FFR,\pi,\xi_t} \pi_t + B_{1,FFR,\pi,\xi_t} \pi_{t-1} + B_{2,FFR,\pi,\xi_t} \pi_{t-2} \\ &\quad + B_{1,FFR,\xi_t} FFR_{t-1} + B_{2,FFR,\xi_t} FFR_{t-2} + \omega_{FFR,t} \end{aligned}$$

where the coefficient $B_{l,i,j,\xi_t}$ gives the impact of variable $i$ to the $l$-lagged variable $j$ under regime $\xi_t$. Note that $-B_{0,FFR,FFR,\xi_t} = 1$. If inflation increases permanently by $\Delta\pi$ we have:

$$\begin{aligned} \Delta FFR &= -B_{0,FFR,\pi,\xi_t} \Delta\pi + B_{1,FFR,\pi,\xi_t} \Delta\pi + B_{2,FFR,\pi,\xi_t} \Delta\pi \\ &\quad + B_{1,FFR,\xi_t} \Delta FFR + B_{2,FFR,\xi_t} \Delta FFR \\ \Delta FFR &= \underbrace{\frac{-B_{0,FFR,\pi,\xi_t} + B_{1,FFR,\pi,\xi_t} + B_{2,FFR,\pi,\xi_t}}{1 - B_{1,FFR,\xi_t} + B_{2,FFR,\xi_t}}}_{LR_{FFR,\pi,\xi_t}} \Delta\pi \end{aligned}$$

To ease notation, in the main text, we define:

$$\psi_{0,x,\xi_t} \equiv -B_{0,FFR,x,\xi_t},\ \psi_{1,x,\xi_t} \equiv B_{1,FFR,x,\xi_t},\ \psi_{2,x,\xi_t} \equiv B_{2,FFR,x,\xi_t}$$

for $x = \pi, \Delta GDP, FFR$. Therefore, we obtain:

$$\begin{aligned} FFR_t &= \psi_{0,\pi,\xi_t}\pi_t + \psi_{1,\pi,\xi_t}\pi_{t-1} + \psi_{2,\pi,\xi_t}\pi_{t-2} \\ &+\psi_{0,\Delta GDP,\xi_t}\Delta GDP_t + \psi_{1,\Delta GDP,\xi_t}\Delta GDP_{t-1} + \psi_{2,\Delta GDP,\xi_t}\Delta GDP_{t-2} \\ &+\psi_{1,FFR,\xi_t}FFR_{t-1} + \psi_{2,FFR,\xi_t}FFR_{t-2} + \omega_{FFR,t} \end{aligned}$$

Note that above $-B_{0,FFR,FFR,\xi_t} = \psi_{0,FFR,\xi_t} = 1$.

## 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 for the mean 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:

$$\mathbb{E}_i\left(Z_t\right) = \overline{\mu}_i = \left(I_n - A_i\right)^{-1} c_i \tag{A13}$$

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 $\mathbb{E}_i\left(Z_t\right)$ over a finite horizon. If there are no further switches, we can then expect the variables to fluctuate around $\mathbb{E}_i\left(Z_t\right)$. Therefore, the conditional steady states for the mean 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.

The conditional steady state for the standard deviation corresponds to the standard deviation for the vector $Z_t$ conditional on being in a particular regime. The conditional standard deviations for the elements in $Z_t$ are computed by taking the square root of the main diagonal elements of the covariance matrix $\mathbb{V}_i\left(Z_t\right)$ obtained imposing that a certain regime is in place forever:

$$\varphi\left(\mathbb{V}_i\left(Z_t\right)\right) = \left(I_{n^2} - A_i \otimes A_i\right)^{-1} \varphi\left(V_{\xi_t} V_{\xi_t}'\right) \tag{A14}$$

where $I_{n^2}$ is an identity matrix with the appropriate size, $\otimes$ denotes the Kronecker product, and the vectorization operator $\varphi\left(X\right)$ takes a matrix $X$ as an input and returns a column vector stacking the columns of the matrix $X$ on top of one another.

## 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 $\pi_t$:

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

where $\pi_t^i = P_0\left(\xi_t = i\right)$ 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 $\pi_t = [\pi_{1,t}, ..., \pi_{m,t}]'$, $\Omega = bdiag\left(A_1, ..., A_m\right) H$, and $C = bdiag\left(c_1, ..., c_m\right)$. Recall that:

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

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

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

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 $q_{t|t}^i = Z_t \pi_t^i$.

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:

$$\underbrace{\theta_{x,t}-\theta_{y,t}+\sum_{j=0}^{\infty}\rho^{j}\mathbb{E}_{t}\left(e_{x,t+1+j}^{*}-e_{y,t+1+j}^{*}\right)}_{\text{Spread in BM ratios corrected for earnings}}=\underbrace{\sum_{j=0}^{\infty}\rho^{j}\mathbb{E}_{t}\left(r_{x,t+1+j}-r_{y,t+1+j}\right)}_{\text{PDV of the expected spread in excess returns}} \tag{A15}$$

The spread in excess returns, $r_{xy,t}\equiv r_{x,t}-r_{y,t}$. Then the right hand side of (A15) can be computed as:

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

Therefore, we have:

$$\widetilde{\theta}_{xy,t}\equiv\underbrace{\widetilde{\theta}_{x,t}-\widetilde{\theta}_{y,t}+\sum_{j=0}^{\infty}\rho^{j}\mathbb{E}_{t}\left(e_{x,t+1+j}^{*}-e_{y,t+1+j}^{*}\right)}_{\text{Spread in BM ratios corrected for earnings}}=1_{r_{xy}}'w\left(I-\rho\Omega\right)^{-1}\left[\Omega q_{t|t}+C\left(I-\rho H\right)^{-1}H\pi_{t|t}\right] \tag{A16}$$

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

Similar formulas are used to compute risk premia for the individual portfolios. The premium for a portfolio $z$ coincides with the present discounted value of its excess returns:

$$\underbrace{premia_{z,t}}_{\text{Premia}}\equiv\underbrace{\sum_{j=0}^{\infty}\rho^{j}\mathbb{E}_{t}\left(r_{z,t+1+j}\right)}_{\text{PDV of excess returns}}=1_{r_{z}}'w\left(I-\rho\Omega\right)^{-1}\left[\Omega q_{t|t}+C\left(I-\rho H\right)^{-1}H\pi_{t|t}\right], \tag{A17}$$

where $1_{r_z}'$ is a vector used to extract the PDV of excess returns from a vector containing the PDV of all variables included in the VAR.

**Regime Average** We also compute the *regime average* value of $\widetilde{\theta}_{xy,t}$. The regime average is defined as:

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

where $\overline{\pi}_{i}=1_{i}$ and $\overline{q}_{i}\equiv[0,...,\overline{\mu}_{i},...,0]$ is a column vector that contains the conditional steady state of for the mean value of $Z_t$ conditional on being in regime $i$, i.e., $\mathbb{E}_{i}\left(Z_{t}\right)=\overline{\mu}_{i}=\left(I_{n}-A_{i}\right)^{-1}c_{i}$, and zero otherwise. Recall that the conditional steady state, $\overline{\mu}_{i}$, is a vector that contains the expected value of $Z_t$ conditional on being in regime $i$. Therefore, the vector captures the values to which the variables of the VAR converge if regime $i$ is in place for a prolonged period of time. Note that $\overline{\widetilde{\theta}}_{xy}^{i}$ is computed by conditioning on the economy being initially at $Z_{t}=\overline{\mu}_{i}$ and in regime $i$, but taking into account that there might be regime changes in the future. Therefore, we can also think about $\overline{\widetilde{\theta}}_{xy}^{i}$ as the expected value of the PDV of excess returns, $\widetilde{\theta}_{xy,t}$, conditioning on being in regime $i$ today and on the variables of the VAR being equal to the conditional steady state mean values for regime $i$. Formally:

$$\overline{\widetilde{\theta}}_{xy}^{i}=\mathbb{E}\left(\widetilde{\theta}_{xy,t}|\xi_{t}=i,Z_{t}=\overline{\mu}_{i}\right). \tag{A18}$$

These are the values around which we expect $\widetilde{\theta}_{xy,t}$ to fluctuate as the VAR variables $Z_t$ fluctuate around $\overline{\mu}_i$.

Similarly, we can compute the *regime average* value of risk premia for an individual portfolio $z$, $premia_{z,t}$:

$$\overline{premia}_z^i \equiv 1'_{r_z} w \left(I - \rho\Omega\right)^{-1} \left[\Omega \overline{q}_i + C \left(I - \rho H\right)^{-1} H \overline{\pi}_i\right]. \tag{A19}$$

Formulas (A16), (A17), (A18), and (A19) are used in the paper to produce Figure 9 and Tables 6 and **??**. For each draw of the VAR parameters from the posterior distribution, we compute the evolution of the PDV of the spread in excess returns over time between high and low return premia portfolios, $\widetilde{\theta}_{xy,t}$ and individual portfolio $premia_{z,t}$, by using (A16) and (A17). Thus, we obtain a posterior distribution for $\widetilde{\theta}_{xy,t}$ and $premia_{z,t}$. The medians of these posterior distributions are reported as the blue solid lines in Figure 9. Similarly, for each draw of the VAR coefficients, we compute $\overline{\widetilde{\theta}}_{xy}^i$ and the difference in the PDV between the two regimes $\overline{\widetilde{\theta}}_{xy}^1 - \overline{\widetilde{\theta}}_{xy}^2$. Thus, we obtain a posterior distribution for $\overline{\widetilde{\theta}}_{xy}^i$ and for the difference $\overline{\widetilde{\theta}}_{xy}^1 - \overline{\widetilde{\theta}}_{xy}^2$. The medians of the distribution of $\overline{\widetilde{\theta}}_{xy}^i$ and $\overline{premia}_z^i$ for $i = 1, 2$, are reported in Figure 9 (red dashed line). Table 6 reports the median and the 68% posterior credible sets both for the distribution of $\overline{\widetilde{\theta}}_{xy}^i$, for $i = 1, 2$, and for the difference in these across regimes, $\overline{\widetilde{\theta}}_{xy}^1 - \overline{\widetilde{\theta}}_{xy}^2$. Finally, the last row of Table 6 reports the percentage of draws for which $\overline{\widetilde{\theta}}_{xy}^1 - \overline{\widetilde{\theta}}_{xy}^2 > 0$ and $\overline{premia}_z^1 - \overline{premia}_z^2 > 0$ as the probability that risk premia are lower in the high asset valuation/low interest rate regime than they are in the low asset valuation/high interest rate regime.

## Variable Selection for VARs to Compute PDV of Risk Premia

We start with a series of fixed regressors that are relevant for predicting market excess returns or the return of the spread portfolios. To limit the size of the MS-VAR, we then use the Akaike information criterion (AIC) to decide whether to include some additional regressors. Specifically, we compute the AIC for the equation(s) that correspond(s) to the return(s) that we are trying to predict. We then choose the specification that minimizes the AIC.

Here are the details:

1. MS-VAR for the Market excess return:

   Fixed regressors (all lagged): Market excess return, inverse valuation ratio based on $cay^{MS}$. The inverse valuation ratio is included because it represents a measure of asset valuation that can predict future stock market returns. Note that given that we are conditioning to the regime sequence obtained when estimating $cay^{MS}$, the intercept for the corresponding equation will adjust in a way to reflect the low frequency breaks identified above.

Possible additional variables to be chosen for the estimation based on the AIC: Value (small) spread (log-difference in the book to market ratio of the small value portfolios and the book to market ratio of the small growth portfolios), Real FFR, term yield spread, four of the five Fama and French factors (SMB, HML, RMW, CMA), *cay* (based on PCE, available on Martin Lettau's website.) Note that we do not include the market excess return from Fama and French (MKTMINRF) as a possible additional regressor because our dependant variable is the excess market return itself. Therefore, this variable is automatically included in the MS-VAR.

Additional regressors selected based on the AIC: Value Spread, and SMB and HML factors from Fama and French.

2. MS-VAR for (a) 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; (b) Value return spread (S1): 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; (c) Value return spread (S2): The difference between the excess return of the size 2 high BM portfolio and the excess return of the small size 2 low BM portfolio.

   Fixed regressors (all lagged): (a) Momentum return spread; (b) Value return spread (S1); (c) Value return spread (S2); (d) 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; (e) Value BM spread (S1): 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; (f) Value BM spread (S2): The difference between the logarithm of the BM ratio of the size quintile 2 high book-market portfolio and the logarithm of the BM ratio of the size 2 low book-market portfolio.

   Possible additional variables to be chosen for the estimation based on the AIC: Real FFR computed as the difference between FFR and Inflation, excess return of small growth portfolio, excess return of small value portfolio, five Fama-French factors (SMB, HML, RMW, CMA, MKTMINRF.)

   Additional regressors selected based on the AIC: Real FFR and excess return of the small value portfolio.