NBER WORKING PAPER SERIES

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

Francesco Bianchi
Martin Lettau
Sydney C. Ludvigson

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

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

Ludvigson acknowledges research support from the C.V. Starr Center for Applied Economics at NYU. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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

ABSTRACT

This paper presents evidence of infrequent shifts, or "breaks," in the mean of the consumption-wealth variable cay_t, an asset market valuation ratio driven by fluctuations in stock market wealth relative to economic fundamentals. Conventional estimates of cay_t, which presume a constant mean, display increasing persistence over the sample. We introduce a Markov-switching version of cay_t that adjusts for infrequent shifts in its mean. The Markov-switching cay_t, denoted cay_t^{MS}, is less persistent and has superior forecasting power for excess stock market returns compared to the conventional estimate. Evidence from a Markov-switching VAR shows that these low frequency swings in post-war asset valuation are strongly associated with low frequency swings in the long-run expected value of the Federal Reserve's primary policy rate, with low expected values for the real federal funds rate associated with high asset valuations, and vice versa. By contrast, there is no evidence that the infrequent shifts to high asset valuations and low policy rates are associated with higher expected economic growth or lower economic uncertainty; indeed the opposite is true.

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

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

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

1 Introduction

In a 2001 published paper, Lettau and Ludvigson (2001) (LL hereafter) introduced an estimated variable shown to have strong forecasting power for U.S. stock returns. It is best thought of as an asset market valuation ratio driven by fluctuations in stock market wealth relative to economic fundamentals. The variable, denoted cay_t, is a stationary linear combination of log consumer spending, c_t, log asset wealth, a_t, and log labor income, y_t, all measured on an aggregate basis in quarterly data. The coefficients of this linear combination may be estimated as parameters of a cointegrating relationship and the variable cay_t is the estimated cointegrating residual of this relationship. Under assumptions described in LL and elaborated on in Lettau and Ludvigson (2010), cay_t may be interpreted as a proxy for the log consumption- aggregate (human and non-human) wealth ratio, and its relationship with future growth rates of a_t (highly correlated with stock market returns in quarterly data) and/or future growth rates of c_t and y_t, may be motivated from an aggregate household budget constraint. This consumption-wealth variable is well described as an asset valuation ratio, or more simply "wealth ratio," because its variability is driven primarily by transitory fluctuations in stock market wealth relative to economic fundamentals (Lettau and Ludvigson (2004)).

In the years since the 2001 paper was published, updated estimates of cay_t have continued to display predictive power for long horizon stock market returns. But over time, the statistical properties of the estimated series appear to have shifted in some fundamental ways. Notably, the measured value of cay_t has become more persistent, resulting forecasting power increasingly concentrated at longer horizons. The rising persistence is also evident from cointegration tests for c_t, a_t, and y_t. Although cay_t has always had a substantial autocorrelation (with a first-order autoregressive coefficient of 0.79 reported in the 2001 paper), theory implies that c_t, a_t, and y_t should be cointegrated, or that cay_t should be covariance stationary. Yet in recent samples it has become difficult, according to some statistical tests, to distinguish cay_t from a unit root process. Similar findings have been documented for other stock market valuation ratios long used as predictor variables for stock returns. These include price-dividend or price-earnings ratios, which are themselves cointegrating residuals.[1] Despite these findings, a literal unit root interpretation for these variables is unappealing because it implies that stock prices or asset values could wander arbitrarily far from measures of fundamental value *indefinitely*.[2] An arguably more appealing interpretation is that there are instead infrequent shifts in certain moments of the stationary distribution that–when not taken into account–make distinguishing

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

[2]Even theories that postulate "bubbles" almost always imply that the bubble will eventually burst, restoring a pre-bubble relationship between prices and fundamentals.

a stationary from a unit root variable difficult in a small sample.

This paper presents evidence of infrequent shifts, or "breaks," in the mean of the consumption-wealth proxy cay_t and introduces a Markov-switching version that adjusts for these shifts. We refer to the regime switching measure as a *Markov-switching cay*, denoted cay_t^{MS}. These infrequent mean shifts generate low frequency fluctuations in asset values relative to fundamentals as measured by cay. Unlike the conventional cay_t, which presumes a constant mean, cay_t^{MS} does not exhibit increasing persistence as estimates are updated over the sample. Moreover, evidence in favor of stationarity for cay_t^{MS} is much stronger in current samples than it is for cay_t. And the estimated persistence of cay_t^{MS}, as well as the point estimates of the cointegrating slope coefficients, are close to values originally reported in LL for cay_t. This suggests that infrequent shifts in the mean of cay_t can help explain why its statistical properties have shifted over time.

After documenting these findings, we first turn our attention to forecasts of the U.S. stock market. We find that the forecasting power of cay_t^{MS} for future stock market returns is superior to that of cay_t, even if no forward-looking data are used in the construction of cay_t^{MS}. This remains true out-of-sample, at least for some horizons. Simulations suggest, however, that this improvement in forecasting power may be harder exploit out-of-sample, unless researchers have a sufficiently large number of observations with which to estimate the Markov-switching parameters. But the simulations also suggest that postwar samples close the size currently available may be large enough to do so going forward. Researchers using cay_t as a predictor variable may wish to consider using both measures to forecast long-horizon stock market returns.[3]

The final section of the paper turns to the key question of what these infrequent mean shifts represent economically. Any estimated statistical relationship is subject to possible structural change as the number of years over which the relationship is measured rises. This may be especially true of cay_t, where the definitions of the embedded variables have changed discretely over time as data collection agencies have altered their measurement criteria for all three series.[4] But structural shifts in the economy are also likely to play a role, as suggested by evidence that other stock market valuation measures have also experienced "breaks" in the mean values of their distributions (e.g., Lettau, Ludvigson, and Wachter (2008); Lettau and Van Nieuwerburgh (2008)). Thus we study what could explain the breaks observed in cay by estimating a Markov-switching macroeconomic VAR (MS-VAR) for output growth, inflation, and the federal funds

[3]Both measures are available on the autors' websites.

[4]As one example, the measurement of a_t (household net worth) was significantly altered in 2013 due to a change in the treatment of defined benefit pension plans. Prior to this time, the plans entered household net worth as the value of the assets held by the funds. From September 2013, these plans were instead accounted for by the present value of promised benefits. Because government pension plans are the majority of these plans and are heavily underfunded, this resulted in a shift up of the whole net worth path. The change can be expected affect the dynamics of cay_t, as the relative economic importance of defined benefit versus defined contribution pension plans has changed over the sample.

rate, allowing the parameters of the VAR to potentially undergo structural breaks during the periods that correspond to the shifts identified in our estimates for cay^{MS}. With this approach, we impose the formerly estimated regime sequence for *cay* on the VAR, but the parameters characterizing the different regimes, as well as the transition matrix, are freely estimated. We find strong evidence of breaks in the long-run expected real federal funds rate that coincide with the breaks in the mean of *cay*, with low wealth ratios (low asset valuations or high *cay*) associated with an expectation of sharply higher values for the real federal funds rate, and high wealth ratios (high asset valuations or low *cay*) associated with expectations of much lower value for the long-run real federal funds rate. By contrast, there is no evidence that these low frequency shifts to high asset valuations and persistently low policy rates are associated with higher expected long-run economic growth or lower economic uncertainty; indeed the opposite is true. These findings therefore run counter to the idea that high asset valuations created by a persistently low interest rate environment are the result of a positive outlook for economic growth, or lower uncertainty about that growth. To the best of our knowledge, these findings are among the first formal statistical evidence that low frequency breaks in asset values relative to economic fundamentals are strongly associated with expectations of the long-run value of the primary policy instrument under direct control of the central monetary authority.

The rest of the paper is organized as follows. The next section of the paper discusses the empirical model and the estimation of a Markov-switching cay_t. Section 3 presents results of this estimation, including evidence of breaks in the mean of *cay*, evidence on the persistence of *cay* once corrected for regime shifts in its mean, and a comparison of the forecasting power of evidence of cay^{MS} and *cay* for U.S. stock market returns. This section also briefly discusses evidence for regime shifts in the mean of a newer measure of cay_t introduced in Lettau and Ludvigson (2015) that uses total personal consumption expenditures as a measure of c_t in place of the usual nondurables and services expenditures employed in the construction of the original cay_t. Section 4 then turns to the question of what could explain the evidence for regime shifts in the mean of the wealth ratio cay_t by estimating a Markov-switching macroeconomic VAR (MS-VAR) for output growth, inflation, and the federal funds rate, allowing the parameters of the VAR to potentially undergo structural breaks during the periods that correspond to the shifts identified in our estimates for cay^{MS}. Section 5 concludes.

2 Empirical Model

The model investigated in this section is as follows. Let $\boldsymbol{z}_t$ be a 3×1 vector of data on c_t, a_t, and y_t, and let $\boldsymbol{Z}_t = \left(\boldsymbol{z}_t, \boldsymbol{z}_{t-1}, ..., \boldsymbol{z}'_{-m}\right)$ be a vector containing all observations obtained through date t. We extend the methodology of LL to allow for possible shifts in the mean in the cointegrating relation between consumption, asset wealth, and income over time. Specifically, we modify the Stock and Watson (1993) dynamic least squares (DLS) regression to allow for shifts in the intercept $\alpha_{\xi_t^\alpha}$:

$$c_t = \alpha_{\xi_t^\alpha} + \beta_a a_t + \beta_y y_t + \sum_{-k}^{k} b_{a,i} \Delta a_{t-i} + \sum_{-k}^{k} b_{y,i} \Delta y_{t-i} + \sigma \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim N(0,1)$. The notation $\alpha_{\xi_t^\alpha}$ indicates that the value of the constant in the above regression depends on the existence of a latent state variable, ξ_t^α. The latent state ξ_t^α is presumed to follow a two-state Markov-switching process with transition matrix $\boldsymbol{H}^\alpha$ so that $\alpha_{\xi_t^\alpha}$ can assume one of two discrete values, α_1 or α_2. The econometric specification (1) permits regime switches only the in the intercept parameter. The Appendix discusses alternative specifications where other parameters are also subject to regime switches.

The parameters of the time-series model for $\boldsymbol{z}_t$ include the cointegrating parameters β_a and β_y, the additional slope coefficients $b_{a,i}$ and $b_{y,i}$ in (1), the two intercept values α_1 and α_2, and the transition probabilities contained in the matrix $\boldsymbol{H}^\alpha$. Collect these parameters into a vector $\boldsymbol{\theta}$.

Absent regime changes, *cay* is defined as:

$$cay_t^{FC} = c_t - (\alpha + \beta_a a_t + \beta_y y_t) - \alpha \tag{2}$$

where the superscript "FC" stands for "fixed coefficients" because the constant α is fixed over time. Notice that when we impose a single regime, the Markov-switching model collapses back to the specification originally used by LL. The variable cay_t^{FC} is the same as that defined in LL where it was denoted cay_t. For the purposes of his paper, we have added the superscript "FC" in order to explicitly distinguish it from the Markov-switching version (3). The parameters $\boldsymbol{\theta}$ of the time-series model for cay_t^{FC} include the cointegrating parameters β_a and β_y, the additional slope coefficients $b_{a,i}$ and $b_{y,i}$ and the single intercept values α.

For the Markov-switching model, the constant $\alpha_{\xi_t^\alpha}$ depends on the regime ξ_t^α. If the sequence $\xi^{\alpha,T} = \{\xi_1^\alpha, ..., \xi_T^\alpha\}$ of regimes in place at each point in time were observed, we could immediately compute cay_t^{MS}. Unfortunately, $\xi^{\alpha,T}$ is generally unobservable and needs to be inferred together with the other parameters of the model. It follows that the two values for the Markov-switching constant $\alpha_{\xi_t^\alpha}$ (α_1 and α_2) must be weighted by their probabilities at each point in time. For this purpose, we consider two estimates of the state probabilities distinguished as *filtered* or *smoothed* probabilities. Let $P(\xi_t^\alpha = i | \boldsymbol{Z}_t; \boldsymbol{\theta}) \equiv \pi_{t|t}^i$ denote the probability that $\xi_t^\alpha = i$ based on data obtained through date t and knowledge of the parameters $\boldsymbol{\theta}$. We refer to these as filtered probabilities. Smoothed probabilities reflect the information about the state at time t that can be extracted from the whole sample: $P(\xi_t^\alpha = i | \boldsymbol{Z}_T; \boldsymbol{\theta}) \equiv \pi_{t|T}^i$. These measures of the state probabilities may be used to construct two versions of a Markov-switching *cay*, based on using either smoothed or filtered probabilities. In both cases, the mean value for *cay* is a probability weighted average of the two intercept coefficients, α_1 and α_2. As a benchmark, we use the smoothed probabilities for our baseline estimate and denote it cay_t^{MS}. When we use filtered probabilities, we use the notation cay_t^{MSfilt}. Thus, cay_t^{MS} is computed one of two ways:

$$cay_t^{MSfilt} = c_t - \left(\sum_{i=1}^{2} \pi_{t|t}^{i} \alpha_i + \beta_a a_t + \beta_y y_t\right). \tag{3}$$

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

We depart in one way from previous work in our estimation of the cointegrating parameters for cay_t^{FC}. Previously, the cointegrating coefficients α, β_a, and β_y are estimated from the static ordinary least squares (OLS) regression

$$c_t = \alpha + \beta_a a_t + \beta_y y_t + \epsilon_t \tag{5}$$

while the DLS regression (1) was used to compute standard errors for the OLS cointegrating coefficients in a manner that adjusts for endogeneity of the regressors. For the Gibbs sampling algorithm described below, it turns out to be numerically more convenient to estimate the cointegrating coefficients in one-step as part of the DLS regression (1). Asymptotically, this will not affect the cointegrating coefficient point estimates.

2.1 Estimation

We use Bayesian methods to estimate regression (1). We first search for the posterior mode using a maximization algorithm. The posterior of the model is obtained by computing the likelihood, as explained in Hamilton (1994), and combining it with the priors. In practice we use a flat prior so our parameter "estimates" are simply those that maximize the likelihood. Nevertheless, from here on our estimation procedure is expressed in Bayesian language to make the use of a Gibbs sampling algorithm sensible, which makes it convenient to do inference and compute additional statistics of interest. The maximized likelihood (posterior mode) is used to obtain initial values for the Gibbs sampling algorithm. Once we have the posterior mode, uncertainty about the parameters and the sequence of regimes can be characterized using the Gibbs sampling algorithm described below.

To simplify notation, we denote the vector containing all variables whose coefficients are allowed to vary over time $x_{M,t}$, while $x_{F,t}$ is used to denote the vector containing all the variables whose coefficients are kept constant. We then obtain:

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

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

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

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

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

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

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

with $V_{\beta,T} = \left(V_{\beta,0}^{-1} + \widetilde{X}_F' \widetilde{X}_F\right)^{-1}$ and $B_{\beta,T} = V_{\beta,T}\left[V_{\beta,0}^{-1} B_{\beta,0} + \widetilde{X}_F' \widetilde{C}_n\right]$, where $\widetilde{C}_n$ 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.

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

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

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

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

where $V_{\alpha_i,T} = \left(V_{\alpha_i,0}^{-1} + \widetilde{X}_{M,i}' \widetilde{X}_{M,i}\right)^{-1}$ and $B_{\alpha_i,T} = V_{\alpha_i,T}\left[V_{\alpha_i,0}^{-1} B_{\alpha_i,0} + \widetilde{X}_{M,i}' \widetilde{C}_{n,i}\right]$ where $\widetilde{C}_{n,i}$ 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: $\alpha_i \sim N(B_{\alpha_i,0}, V_{\alpha_i,0})$ 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.

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

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

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

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

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: $\sigma \sim IG(Q_0, v_0)$. With flat priors, we have $Q_0 = 0$ and $v_0 = 0$.

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

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

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

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

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

For each draw of the parameters β_r, $\alpha_{\xi_t^\alpha,r}$, σ_r, and H_r^α we can then compute an estimate for cay_t^{MS} using the filtered, $\pi_{t|t}$, or smoothed probabilities, $\pi_{t|T}$, of the regimes conditional on the model parameters. 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 $\boldsymbol{\theta} = \left\{\beta, \alpha_{\xi_t^\alpha}, \sigma, H^\alpha\right\}$. They can be obtained using the following recursive algorithm:

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

where η_t is a vector whose j-th element contains the conditional density $p(c_t|\xi_t^\alpha = j, x_{M,t}, x_{F,t}; \beta, \alpha_{\xi_t^\alpha}, \sigma)$, the symbol $\odot$ denotes element by element multiplication, and $\mathbf{1}$ is a vector with all elements equal to 1. To initialize the recursive calculation we need an assumption on the distribution of ξ_0^α. We assume that the two regimes have equal probabilities: $p(\xi_0^\alpha = 1) = .5 = p(\xi_0^\alpha = 2)$.

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

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

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

[5] 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).

3 Results

We estimate the Markov-switching cointegrating relation described by (1) over the sample 1952:Q1-2013:Q3 using six leads and lags. Table 1 reports the parameter estimates, while Figure 1 reports the probability of Regime 1 for the Markov-switching intercept $\alpha_{\xi_t^\alpha}$ based on the posterior mode parameter estimates. The 90% credible sets are obtained making 2,000,000 draws from the posterior using the Gibbs sampling algorithm described above. One in every one hundred draws is retained. We check convergence using the methods suggested by Geweke (1992) and Raftery and Lewis (1992).[6]

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

Table 1 provides estimates of the *difference* between the two means and its distribution. The difference is positive and statistically significant, as exemplified by the third row of Table 1, which shows that a 95% credible set only contains non-zero and positive values for this difference.[7] The two regimes turn out to be very persistent and this is reflected in the estimates for the diagonal elements of the transition matrix H^α, also reported in Table 1.

The mode values for the other cointegration parameters are $\beta_a = 0.26$ and $\beta_y = 0.62$. These values are comparable with those originally obtained by LL using a fixed coefficient regression ($\beta_a = 0.31$ and $\beta_y = 0.59$). By contrast, Table 2 reports the parameter estimates for the fixed coefficient cointegrating relation over the extended sample used in this paper, where $\beta_a = 0.12$ and $\beta_y = 0.78$. Therefore, in our current sample, the fixed coefficient parameter estimates differ substantially from those reported in 2001. Bearing in mind that deviations from the cointegrating relation are the result of persistent but transitory movements in a_t rather than c_t or y_t (Lettau and Ludvigson (2004), Lettau and Ludvigson (2013)), these results suggest that

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

[7]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.

the fixed-coefficient estimates of cay_t attempted to "compensate" for increasingly persistent deviations in a_t from its cointegrating relation with c_t and y_t, by progressively reducing the weight on a_t and increasing the weight on y_t. The instability in these point estimates is largely eliminated by allowing for discrete shifts in the mean of *cay*.

To give a visual impression of these regimes over time, Figure 2 plots $cay_t^{MS} + \sum_{i=1}^{2} \pi_{t|T}^{i} \alpha_i$ over time, which is the estimated Markov-switching *cay* from (4) *inclusive* of the intercept. 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 *cay* 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.

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

Figure 3 plots the fixed coefficient cay_t^{FC} and the Markov-switching cay_t^{MS} as defined in (2) and (4), respectively.[8] (Unlike Figure 2, these values subtract the estimated mean or probability-weighted mean, respectively.) The two vertical bars mark the beginning and the end of the time span during which the high mean regime was most likely to be in place. As Figure 3 shows, cay_t^{FC} exhibits persistent deviations from zero, especially during the period starting around 1980 and ending in the early 2000s. This period roughly coincides with the occurrence of Regime 1 when allowing for a Markov-switching constant. By contrast, cay_t^{MS} does not exhibit such persistent deviations from its demeaned value of zero. The persistent deviations are instead captured by low-frequency regime changes in the constant of the cointegrating relation.

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

Several other tests are employed to assess the degree of persistence in cay_t^{MS} as compared

[8] In using the DLS regression (1) 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. For estimates of cay_t^{MS}, we need the filtered and smoothed probabilities for these six lead and lags periods as well. To obtain the estimated probabilities for these points in the sample, we fix the parameters to the values estimated with the 6 leads and 6 lags and proceed as follows. Filtered probabilities for the first observation are obtained using a regression that includes 6 leads but no lags. Filtered probabilities for the second observation are obtained using a regression with 6 leads and 1 lag; estimates for the third observation are obtained using a regression with 6 leads and 2 lags, and so on until we reach the standard regression with 6 leads and 6 lags. To obtain filtered probabilities for the last 6 observations, we proceed the other way around: The last observation estimates are based on a regression that has 6 lags and 0 leads, the next to last is based on a regression that has 6 lags and 1 lead and so on. Smoothed probabilities are then computed with standard methods as they only involve the filtered probabilities and the transition matrix $\mathbf{H}^{\alpha}$.

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

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

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

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

As a third way to evaluate the persistence of in cay_t^{MS} versus cay_t^{FC}, we consider a parameterization from a fractionally integrated model to formalize the heteroskedasticity shown in 4. We therefore assume $(1-L)^d cay_t = u_t$, where L is the lag operator and u_t is an $I(0)$ process. If cay_t is $I(0)$, then $d = 0$. If cay_t has a unit root, then $d = 1$. Non-integer values of $d > 0$ are fractionally integrated series that are more persistent than $I(0)$ but less persistent than $I(1)$. Figure 5 shows the estimated log likelihoods for $(1-L)^d\, cay_t^{MS}$ and $(1-L)^d\, cay_t^{FC}$ as a function of d. For cay_t^{MS}, the likelihood peaks at $d = 0$, while for cay_t^{FC}, the likelihood rises with $d > 0$ and peaks near $d = 1.2$.[9]

Although the statistical tests just considered imply that cay_t^{FC} is sufficiently persistent that it can be difficult to distinguish from a unit root process, it does not follow that cay_t^{FC} actually has a unit root. Tests of the null of no cointegration are known to have low power against

[9]Please refer to Appendix 5 for details about the estimation of the fractionally integrated model.

the cointegration alternative when deviations from the common trend are stationary but highly persistent. For this reason, Park (1990), Park (1992), Han and Ogaki (1997), and Ogaki and Park (1997) developed tests for the null of *cointegration*, rather than no cointegration, which they argue are more appropriate when theory suggests the variables should be cointegrated. These tests, as they apply to *cay* specifically, are described in detail in the online appendix to (Lettau and Ludvigson (2013)) available on the authors' websites. Updated output from the Ogaki and Park (1997) test for the null of cointegration for cay_t^{FC} is provided in Table 4. As in earlier samples, this test continues to show no evidence against the null of cointegration for cay_t^{FC}, lending support to the hypothesis that the standard *cay* is stationary even if it is sufficiently persistent so as to make it difficult to distinguish from a non-stationary variable in our sample. But even if cay_t^{FC} is simply highly persistent but ultimately stationary, the resulting low frequency deviations from a fixed mean raise issues for forecasting. We turn to these forecasting implications next.

3.2 Forecasts of Excess Stock Market Returns

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

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

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

Table 7 reports mean-square forecast errors (MSEs) from out-of-sample forecasts. The forecasting relation is estimated in an initial period using data available from 1952:Q1 through 1980:Q4. Forecasts over the next h quarters are computed and forecast errors stored. The forecasting relation is then reestimated in rolling subsamples moving forward, (i.e., over the period 1952:Q1 through 1981:Q1), and forecasts and forecast errors are computed over the next h periods. This process is repeated until the end of the sample. Table 7 reports MSEs for several forecasting regressions. To form a basis for comparison, the first row reports results using nothing more than an estimated constant as a predictor variable, while the second row uses the lagged log excess stock market return as a predictor. The next three rows report results using cay_t^{FC}, cay_t^{MSfilt}, and cay_t^{MS} as univariate predictors. These versions of cay are all estimated using the full historical sample, as explained above. The last two rows report results using cay_t^{FCrec} and cay_t^{MSrec} as univariate predictors. These versions of cay are all estimated using only data up to and including date t, as explained above.

Some researchers have argued that many predictor variables for stock market returns have difficulty beating the sample mean of stock returns in out-of-sample tests (e.g., Goyal and Welch (2003); Goyal and Welch (2008)). The first row of Table 7 shows this is not the case here: all versions of cay have substantially lower out-of-sample MSEs than a forecasting model that uses only the (constant) sample mean of excess returns as a predictor, and even the recursively estimated versions have MSEs that are almost 70% smaller than those of the sample mean model. Table 7 shows that all versions of *cay* also have lower lower MSEs than a simple autoregressive forecasting model. Among those versions that are estimated using the full sample, the two Markov-switching versions, cay_t^{MSfilt}, and cay_t^{MS}, are much better predictors than the fixed-mean version cay_t^{FC}, having MSEs that are almost 50% smaller for 16-quarter return forecasts. The recursively estimated versions cay_t^{FCrec} and cay_t^{MSrec} have about the same predictive power over most horizons, although the Markov-switching *cay* offers a slight improvement over the fixed-mean *cay* at the longest (16 quarter) horizon. Because these recursive versions are estimated over short subsamples, the estimates of parameters are much noisier than they are for the full-sample versions, so it is not surprising that they have higher MSEs. For this very reason, it is likewise encouraging that cay_t^{MSrec} preforms as well (and slightly better at long horizons) as cay_t^{FCrec}, given that the former has many more parameters that require estimation over short subsamples of our quarterly dataset. Postwar samples of the size currently available are, however, much larger than the repeated subsamples used to construct the recursive estimates for this exercise. Going forward, such samples should provide less noisy estimates of *cay* parameters. Researchers using cay_t as a predictor variable may wish to consider both measures as predictors of long-horizon stock market returns.[10]

3.3 *cay* Measured With Total PCE

We now explore evidence for regime shifts in the mean of a newer measure of cay_t introduced in Lettau and Ludvigson (2015) that uses total personal consumption expenditure (PCE) as a measure of c_t in place of nondurables and services (NDS) expenditures employed in the construction of the original cay_t and used for all the results above. The common practice of using NDS expenditures as a proxy total consumption is well understood to omit a component of total consumption, namely the service flow from the stock of household durable goods. This service flow is unobserved and no reliable measures exists for the durables stock as a whole. Expenditures on durable goods, which are observed, represent investment (replacements and additions to the durable goods capital stock), rather than a flow of consumption. If NDS expenditures were a stable fraction of total flow consumption, the two would be approximately proportional and the dynamic relationship between the log values of either variable with log labor income, y_t, and log asset wealth, a_t, would be the same. Such stability would justify the use of NDS in place of the unobserved total flow consumption. Looking at data over the last 40 years, however, it has become clear that such stability is likely to be illusory (Figure (6)). The data exhibit a sharp downward trend in the ratio of NDS to total PCE over last 40 years. Although PCE is not equivalent to total flow consumption, any plausible model of the

[10]Both series will be updated and available on the authors' websites.

service flow from the durables stock would imply a similar trend in the ratio of NDS to total consumption.

The clear secular trend in NDS expenditures relative to total PCE, whatever its cause, means that it is no longer tenable to ignore durables expenditures when constructing a measure of cay_t with fixed coefficients. Such trends would introduce a non-stationarity into the cointegrating residual of a fixed coefficient cay_t. One way that this might be addressed is to model regime switches in the parameters, as we have done above. Allowing for low frequency regime changes in the appropriate parameters could correct for the nonstationarity induced by low frequency shifts in the NDS/PCE ratio. The findings above are suggestive that a regime-switching mean partly corrects for this source of nonstationarity.

We have also experimented with using crude proxies of the service flow from the household durable stock, but lacking compelling data on key features of this flow, such measures appear unsatisfactory. Instead, in the 2015 update of cay_t, we have simply replaced NDS expenditures with total PCE as our measure of "c_t" in the fixed coefficient measure of cay_t. Lettau and Ludvigson (2015) argue that, under empirically plausible assumptions, the log of total consumption is cointegrated with the log of PCE, with cointegrating vector $(1, -1)'$. It follows that we can use total PCE in place of unobserved total consumption, and still obtain a valid cointegrating relation with a_t and y_t.

We redid our analysis above for this newer measure of cay_t that uses total PCE. In order to conserve space, we do not report all the tables reported above corresponding to the NDS measure, and instead focus only on the forecasting performance of the MS versions compared to the fixed mean versions. Overall, the results show that the PCE measure of cay_t is slightly less persistent even with a fixed mean than is the NDS measure. It has a first-order autoregressive coefficient of 0.91 instead of 0.94, and estimates of the fractionally integrated model imply that it has a root that is just below unity (indicating fractional integration, rather than unit root). Moreover, the Ogaki and Park tests for the null of cointegration typically have sample test statistics that indicate more evidence of cointegration.

This may explain why we find that allowing for breaks in the mean of the PCE cay doesn't result in a large impact on its stationarity. The estimated PCE cay^{MS} is less persistent than the fixed coefficient version, but the cosine transformations continue to exhibit a peak at low frequencies and the log likelihood values for $(1-L)^d cay$ continue to peak in the fractionally integrated ranges. The estimated regimes are less persistent than those for the NDS cay^{MS}, so there is more switching between the two regimes over the sample. The large infrequent shifts in mean that render the NDS version of cay^{MS} much more clearly stationary than the fixed coefficient counterpart are simply not present in the PCE version.

Table 6 presents long-horizon forecasts analogous to those in Table 5, while the bottom panel of Table 7 presents out-of-sample results. Taken together, the results show that the fixed coefficient PCE *cay* has stronger forecasting power than the fixed coefficient NDS *cay*. But the MS version of the latter performs better than the MS of the former, and has the strongest forecasting power of all the measures. This is true both in and out-of-sample. These measures are also the least persistent. These findings echo those in previous work studying the dividend-price ratio. Lettau and Van Nieuwerburgh (2008) remove the non-stationary component of the dividend-price ratio by estimating a structural break model of its mean. Once this ratio is adjusted for structural shifts in its mean, the resulting adjusted process is far less persistent than the original series and exhibits stronger forecasting power for future returns than the unadjusted series.

The finding that the fixed coefficient PCE version of *cay* is somewhat less persistent than the fixed coefficient NDS version suggests that a part of the persistence in the latter may be attributable to factors related to the measurement of consumption. A possible contributing factor to the increasing divergence between NDS and PCE expenditures is the sharply rising income and (to a lesser extent) wealth inequality over this same period, which has increased the income and wealth of households at the very top of these distributions relative to those elsewhere in the distribution. Although we lack high quality measures of the consumption for the wealthy, it stands to reason that consumption inequality has likely grown in tandem with income and wealth inequality. If luxury goods are disproportionately durable goods, then NDS expenditures would omit and increasingly important component of aggregate consumption driven by the consumption of the wealthy that should be cointegrated with aggregate wealth and labor income. Hence using total PCE in *cay* could help correct for some of these biases.

At the same time, the PCE version of *cay* is still quite persistent, regardless of whether the mean is allowed to switch, suggesting that these measurement issues may not be the whole story behind its persistence. Because the PCE measures with fixed mean are less persistent than the NDS measure, while the MS PCE measure is much more persistent than the the MS NDS measure, there seems to be a clear ranking in terms of forecasting power: the MS NDS measures perform best, the fixed coefficient PCE measure performs second best, and the fixed coefficient NDS measure performs least well. Data on the first two measures will be posted regularly on the authors' websites.

4 Explaining Breaks in *cay*: Monetary Policy Regimes

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

Specifically, we consider the following multivariate model:

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

where Z_t is an $n \times 1$ vector of variables, c_{ξ_t} is an $n \times 1$ vector of constants, A_{l,ξ_t} for $l = 1, 2$ is an $n \times n$ matrix of coefficients, $V_{\xi_t} V'_{\xi_t}$ is an $n \times n$ covariance matrix for the $n \times 1$ vector of shocks ε_t. The process ξ_t controls the regime that is in place at time t and assumes two values, 1 and 2, based on the regime sequence identified in our estimates for cay^{MS}.

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

Figure 7 and Figure 8 report the conditional means and conditional standard deviations for the three observables in the VAR and the implied real interest rate at different horizons. The conditional means and the conditional standard deviations for a variable z_t correspond to $\mathbb{E}_t(z_{t+s})$ and $sd_t(z_{t+s}) = \sqrt{\mathbb{V}_t(z_{t+s})} = \sqrt{\mathbb{E}_t[z_{t+s} - \mathbb{E}_t(z_{t+s})]^2}$, where both statistics are computed by taking into account the possibility of regime changes. (Details about how these are calculated can be found in Appendix.) The real interest rate (RIR) is defined as the difference between the FFR and one-step-ahead inflation expectations. Inflation expectations are, in turn, computed based on the VAR estimates. Therefore, the real interest rate is not included directly in the VAR, but derived ex-post based on the VAR estimates.

Figure 7 shows striking evidence of structural change in the expected long-run RIR that coincide with the regime sequence estimated for the mean of *cay*. The occurrence of regime 1, a period of low equity valuation (high *cay*) in the middle subsample from 1976:Q2-2001:Q2, coincides with an expectation of sharply higher values for the real federal funds rate, while the periods of high equity valuation at the beginning (1955:Q3-1976:Q1) and end (2001:Q3-2013:Q3) of our sample coincide with expectations of much lower real interest rates. The differences across subsamples are strongly statistically significant. Note that, because the MS-VAR parameters are freely estimated, the estimation could have found no evidence of structural change in the expected real interest rate across these subsamples and/or that changes occur in variables other than the expected real interest rate.

Figure 7 also shows that the estimated regime shifts in the expected future RIR show up most prominently in the expectations for the real policy rate five to ten years ahead. This finding underscores the extent to which low frequency shifts in the mean of *cay* coincide with expectations of a persistent low or high interest rate environment, rather than transitory movements in these rates. Moreover, the breaks in the expected real interest rate five to ten years ahead appear mostly attributable to breaks in the conditional mean of the *nominal* interest rate, which the Federal Reserve directly influences. Of course the Federal Reserve may also have considerable influence over expected inflation. But movements in expected inflation do not line up as well with the regime sequence for breaks in the mean of *cay* as do movements in the expected nominal interest rate: in the first subperiod, corresponding to the first instance of regime 1, expected inflation was low and then high, while in the second subperiod, corresponding to regime 2, inflation was high and then low, where it remained throughout the entire span of the third subperiod corresponding to the second instance of regime 2 at the end of our sample. To the best of our knowledge, these findings provide among the first formal statistical evidence that low frequency shifts in asset values relative to economic fundamentals are strongly associated with expectations of the long-run value of a policy instrument under direct control of the central monetary authority.

Why are high asset valuations, as measured by the breaks in the mean of *cay*, associated with low expected long-run policy rates, and vice versa? Some theories of rational bubbles suggest that higher policy rates can lead to *higher* asset valuations (e.g., Galí (2014)). But the evidence here is inconsistent with this story, since high wealth ratios are associated with *low* policy rates rather than high. An alternative explanation, consistent with the evidence here, is that any change in the expected short-term real interest rate will always have some effect on asset values because it changes the "fundamental" value of the asset. If prices are inflexible and the Federal Reserve reduces the nominal interest rate, changes in monetary policy may reduce the rate at which investor's discount future payouts by reducing the real "risk-free" rate component of the discount rate, thereby increasing asset values. Asset valuations could be further increased if the risk premium component of the discount rate falls simultaneously with the risk-free rate because investors' willingness to tolerate risk is for some reason inversely related to the long-run expected value of the Federal Reserve's core policy instrument, as in some "reaching for yield" stories.

Alternatively and/or in addition, high wealth ratios could be associated with low expected long-run policy rates because the latter are expected to generate either faster long-run economic growth, or lower uncertainty about that growth. Conversely, regimes characterized by low wealth ratios and high expected policy rates could be explained by lower expectations for long-run growth and/or higher uncertainty about that growth. Figure 7, however, provides no evidence that the low frequency shifts to high asset valuation regimes are associated with higher expected economic growth, or vice versa; indeed the opposite is true. The high asset valuation regime (low *cay* regime 2) at the beginning and end of our sample is associated with *lower* expected GDP growth 10 years ahead than the low asset valuation regime in the middle of the sample. Nor is there a clear pattern in long run expectations of inflation that could explain the corresponding wealth ratio regimes. As mentioned, expected inflation was high in the first subsample (regime 2) but lower in both the middle and ending subsamples (regimes 1 and 2, respectively).

In theory, higher asset valuations could be the result of lower expected economic uncertainty (e.g., Lettau, Ludvigson, and Wachter (2008)). Figure 8 shows the estimated conditional standard deviations for each variable in the MS-VAR across our sample. Note that the conditional standard deviation represents a statistical measure of uncertainty. The result in Figure 8 shows that for GDP growth and inflation, uncertainty is *higher* rather than lower in subperiods of high equity valuation (regime 2) as compared to the subperiod of low equity valuation (regime 1), but the opposite is true for the nominal and real federal funds rate. Thus infrequent shifts to high mean wealth ratios cannot be readily explained by lower economic uncertainty, nor can it be explained by lower inflation uncertainty. Therefore, we can conclude that the high asset valuation regime 2 is characterized by higher uncertainty for real activity and inflation, but lower uncertainty about the Federal Reserve's policy instrument, while the converse is true for regime 1. This result is consistent with a more active role of the Federal Reserve in stabilizing inflation and real activity. As the Federal Reserve is expected to respond more aggressively by raising interest rates to counter higher inflation and/or a lower output gap, macroeconomic volatility is reduced, whereas the volatility of the FFR *can* increase.

Table 8 reports both means and standard deviations for the real interest rate and GDP growth, conditional on staying in each of the two regimes. These conditional steady state values are the means and volatilities that we would expect to observe if one regime were to prevail for a prolonged period of time.[11] Note that these estimates do not depend on the estimated transition matrix. These statistics corroborate the non-steady state evidence presented above where the possibility of a regime shift is incorporated into expectations: the two regimes present a clear

[11] Formally, the conditional means and standard deviations are the values to which $\mathbb{E}_t\left(rir_{t+s}\right)$ and $sd_t\left(z_{t+s}\right) = \sqrt{\mathbb{V}_t\left(rir_{t+s}\right)}$ would converge if one of the two regimes were in place forever as the horizon s goes to infinity.

difference for the mean and volatility of the real interest rate. The high asset valuation regime 2 is characterized by sharply lower expected real policy rates and lower uncertainty about those rates, while the opposite is true for the low asset valuation regime. By contrast, the high asset valuation regime 2 is characterized by lower expected economic growth and higher economic uncertainty.

Figure 9 gives a visual impression of the key result. The figure plots the "wealth ratio" (the inverse of cay^{MS} without removing the Markov-switching constant), along with the ten-year-ahead expected real federal funds rate, on separate scales. The red dashed line in the figure shows the most likely value of the unconditional mean of the wealth ratio in each regime (given by the inverse of the regime-probability weighted average of α_1 and α_2). The mean shows clear regime shifts in wealth ratios that move from high to low to high over the sample, coinciding with a low then high then low expected long-run real federal funds rate. Note that the regime shifts in the expected federal funds rate are large, ranging from about 1% in the low expected interest rate regimes to 3% in the high expected interest rate regime.

These results capture three distinct periods of post-WWII US economic history. The first manifestation of regime 2 is in the subperiod from 1952:Q1-1976:Q1 and coincides with the run-up of inflation in the 1960s and 1970s, accommodative monetary policy, and low real interest rates. Economists have provided several possible explanations for why monetary policy failed to react aggressively to inflation during those years. However, they generally tend to agree that this was a period of high uncertainty and possibly passive monetary policy (Clarida, Gali, and Gertler (2000); Lubik and Schorfheide (2004); Sims and Zha (2006); Bianchi (2013)). The occurrence of Regime 1, in the middle subperiod from 1976:Q2-2001:Q2, proceeds by three years Volcker's appointment as Chairman of the Federal Reserve and roughly coincides with the Volcker disinflation that followed and the Great Moderation. The first attempts to bring inflation down started in the late 1970s, whereas the beginning of Great Moderation is generally placed in the mid-1980s. Macroeconomists interested in the Great Inflation tend to identify the change in the anti-inflationary stance of the Federal Reserve with the appointment of Volcker in August 1979. However, Sims and Zha (2006) estimate a structural MS-VAR and find a change in the conduct of monetary policy from less to more active toward the end of 1977, in line with our results. Real interest rates increased significantly during the Volcker disinflation and they remained higher than in the 1970s for a prolonged period of time. In part this was probably due to the fact that the Federal Reserve had to rebuild credibility for low and stable inflation.

At the same time, the economy experienced a substantial reduction in volatility (McConnell and Perez-Quiros (2000); Stock and Watson (2002)). Finally, the second occurrence of Regime 2 in the subperiod 2001:Q3-2013:Q3 starts with the end of the information technology (IT) boom and the start of the Federal Reserve's accommodative response to the recession that followed. Economists have identified the end of the Great Moderation with the 2008 recession, consistent with the estimated break patterns in Figure 8 for GDP growth uncertainty. At the same time, some have argued that monetary policy underwent a regime shift after the end of the IT boom (Campbell, Pflueger, and Viceira (2014)) and/or that interest rates were held "too low for too long"(Taylor (2007)) in response to the IT bust and the aftermath of 9/11. Asset values quickly recovered in 2002, and after a brief but dramatic decline in the financial crisis of 2007-2009, equity valuations resumed their upward march in 2009. This period of high equity valuations persists today with the zero lower bound associated with positive rates of inflation. Our estimates characterize this third subperiod as a return to a period of prolonged low real interest rates, i.e. regime 2.

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

5 Conclusion

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

At the same time, we find no evidence that the estimated structural shifts to high asset valuation regimes and persistently low policy rates are associated with optimism about the future in the form of expectations for stronger long-run economic growth, or lower uncertainty about that growth. Instead, the results suggest that infrequent regime shifts to high asset valuations may be driven by persistent shifts in the stance of monetary policy that merely reduce the rate at which investors discount assets, without engendering favorable expectations for real economic growth, inflation, or uncertainty.

Appendix

Data Appendix

This appendix describes the data used in this study.

CONSUMPTION

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

LABOR INCOME

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

POPULATION

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

WEALTH

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

STOCK PRICE, RETURN, DIVIDENDS

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

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

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 $\ln\left(P_t^{Ax}/D_t^A\right)$.

PRICE DEFLATOR

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

Estimation of Fractionally Integrated Models

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

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

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

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

$$p_{X^S}(x^s) = \frac{1}{2}\Gamma(k/2)\, \pi^{-k/2} \left|\Sigma_X\right|^{-1/2} \left(x^{s\prime} \Sigma_X^{-1} x^s\right)^{-q/2} \tag{7}$$

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

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

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

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

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

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

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

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

Finally, we evaluate (7) to obtain the likelihood for the different values of d given that $\Sigma_X(d)$ is now a function of the parameter d of the fractionally integrated model.

Robustness: Markov-Switching in Other Parameters

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

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

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

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

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

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

Additional Statistical Results

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

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

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

MS-VAR estimation

In this appendix we provide details on the estimation of the MS-VAR 6. As explained above, we take the regime sequence as given based on our estimates for the breaks in cay^{MS}. This means that we only have to estimate the transition matrix and the parameters of the 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. Conditional on each regime, we can rewrite the VAR as

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

where $\Sigma_{\xi_t} = V_{\xi_t} V_{\xi_t}'$. If we specify a Normal-Wishart prior for β_{ξ_t} and V_{ξ_t}:

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

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

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

where

$$\begin{aligned} v_T &= T + v_0, \quad N_T = X'X + N_0 \\ B_T &= N_T^{-1}\left(N_0 B_0 + X'X\widehat{B}_{MLE}\right) \\ S_T &= \frac{v_0}{v_T}S_0 + \frac{T}{v_T}\widehat{\Sigma}_{MLE} + \frac{1}{v_T}\left(\widehat{B}_{MLE} - \widehat{B}_0\right)' N_0 N_T^{-1} X'X\left(\widehat{B}_{MLE} - \widehat{B}_0\right) \\ \widehat{B}_{MLE} &= \left(X'X\right)^{-1}\left(X'Y\right),\ \widehat{\Sigma}_{MLE} = \frac{1}{T}\left(Y - X\widehat{B}_{MLE}\right)'\left(Y - X\widehat{B}_{MLE}\right) \end{aligned}$$

We choose flat priors ($v_0 = 0$, $N_0 = 0$) so the expressions above coincide with the maximum likelihood estimates (MLE):

$$v_T = T,\ N_T = X'X,\ B_T = \widehat{B}_{MLE},\ S_T = \widehat{\Sigma}_{MLE}.$$

Finally, the posterior of the transition matrix H of the MS-VAR only depends on the regime sequence $\xi^{\alpha,T}$ estimated for cay^{MS}. Given the draws for the MS state variables $\xi^{\alpha,T}$, 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.[12] For each column of H the posterior distribution is given by:

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

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

[12] 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).

Expectations

In this appendix we explain how expectations and uncertainty are computed for the MS-VAR. More details can be found in Bianchi (2016). Consider the following multivariate Markov-switching model:

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t}\varepsilon_t,\ \varepsilon_t \sim N(0, I) \tag{8}$$

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 any VAR with $l > 1$ lags can be rewritten as above by using the companion form. Let n be the number of variables in the VAR of the previous Appendix section. Define the $mn \times 1$ column vector q_t as:

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

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

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

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

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

where the matrix $w = [I_n, ..., I_n]$ 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 (8) and define* $q_t^i = \mathbb{E}_0\left(Z_t 1_{\xi_t=i}\right)$ *for* $i = 1...m$. *Then* $q_t^j = c_j\pi_t^j + \sum_{i=1}^m A_j q_{t-1}^i h_{ji}$.

Using this result, we can write the law of motion of q_t as:

$$q_t = C\pi_t + \Omega q_{t-1} \tag{9}$$

$$\pi_t = H\pi_{t-1} \tag{10}$$

with $\pi_t = [\pi_t^1, ..., \pi_t^m]'$, $\Omega = bdiag(A_1, ..., A_m)(H \otimes I_n)$, and $C = bdiag(c_1, ..., c_m)$, where $\otimes$ 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. The central insight of (9)-(10) consists of the fact that while Z_t is not Markov, q_t is. It is then straightforward to compute expectations conditional on the information available at a particular point in time. Suppose we are interested in $\mu_{t+s|t} = \mathbb{E}_t(Z_{t+s})$, 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 $q^i_{t+s|t} = \mathbb{E}_t\left(Z_{t+s}1_{\xi_{t+s}=i}\right)$, we have $\mathbb{E}_t\left(Z_{t+s}\right) = wq_{t+s|t}$.

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

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

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

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

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

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

Proposition 2 *Consider a Markov-switching model whose law of motion can be described by (8) and define* $Q^i_t = \varphi\left[\mathbb{E}_0\left(Z_tZ_t'1_{\xi_t=i}\right)\right]$, $q^i_t = \mathbb{E}_0\left[Z_t1_{\xi_t=i}\right]$, *and* $\pi^i_t = P_0\left(\xi_t = i\right)$, *for* $i = 1...m$. *Then* $Q^j_t = \left[\widehat{cc}_j + \widehat{VV}_j\varphi\left[I_k\right]\right]\pi^j_t + \sum_{i=1}^m\left[\widehat{AA}_jQ^i_{t-1} + \widehat{DAC}_jq^i_{t-1}\right]h_{ji}$, *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* $\widehat{DAC}_j = \left(A_j \otimes c_j\right) + \left(c_j \otimes A_j\right)$.

Using matrix algebra we obtain:

$$Q_t = \Xi Q_{t-1} + \widehat{DAC}q_{t-1} + \widehat{V}c\pi_t \tag{11}$$

$$q_t = C\pi_t + \Omega q_{t-1},\ \pi_t = H\pi_{t-1}. \tag{12}$$

where

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

Even in this case, the central insight consists of the fact that while Z_t is not Markov, Q_t is. It is then straightforward to compute the evolution of second moments conditional on the information available at a particular point in time. Suppose we are interested in $\mathbb{E}_t\left(Z_{t+s}Z_{t+s}'\right)$, i.e. the second moment of the vector Z_{t+s} conditional on the information available at time t. If we define:

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

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

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

$$\varphi\left[\mathbb{V}_t\left(Z_{t+s}\right)\right] = M_{t+s|t} - \varphi\left[\mu_{t+s|t}\mu_{t+s|t}'\right]. \tag{13}$$

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

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

References

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

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.

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

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

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.

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

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

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

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

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

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

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

Kim, C.-J., and C. R. Nelson (1999): *State-Space Models with Regime Switching*. The MIT Press, Cambridge, Massachusetts.

Lettau, M., and S. C. Ludvigson (2001): “Consumption, Aggregate Wealth and Expected Stock Returns,” *Journal of Finance*, 56(3), 815–849.

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

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

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

Lettau, M., and S. C. Ludvigson (2015): “Changes in the Measurement of Consumption in *cay*,” http://faculty.haas.berkeley.edu/lettau/data/Revision_PCE_2015.pdf.

Lettau, M., S. C. Ludvigson, and J. A. Wachter (2008): “The Declining Equity Premium: What Role Does Macroeconomic Risk Play?,” *Review of Financial Studies*, 21(4), 1653–1687.

Lettau, M., and S. Van Nieuwerburgh (2008): “Reconciling the Return Predictability Evidence: In-Sample Forecasts, Out-of-Sample Forecasts, and Parameter Instability,” *Review of Financial Studies*, 21(4), 1607–1652.

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

Lubik, T. A., and F. Schorfheide (2004): “Testing for indeterminacy: an application to US monetary policy,” *American Economic Review*, pp. 190–217.

McConnell, M. M., and G. Perez-Quiros (2000): “Output Fluctuations in the United States: What Has Changed Since the early 1980s?,” *American Economic Review*, 90(5), 1464–1476.

Muller, U. K., and M. W. Watson (2008): “Testing Models of Low Frequency Variability,” *Econometrica*, 76, 155–74.

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

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

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

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

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

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

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

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

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

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

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

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

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

Tables and Figures

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

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

Parameter Estimates: cay^{FC}

α	β_a	β_y
0.8706 (0.0345)	0.1246 (0.0150)	0.7815 (0.0168)

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

Cointegration Tests

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

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

Canonical Cointegrating Regression Results

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

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

Long Horizon Forecasting Regressions: Stock Returns

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

	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]
cay^{MSfilt}	**1.54**	**6.38**	**11.60**	**13.56**	**13.61**
	(4.07)	(5.22)	(6.53)	(6.03)	(6.18)
	[0.04]	[0.18]	[0.35]	[0.37]	[0.34]
cay^{MS}	**1.49**	**6.83**	**11.88**	**13.79**	**13.78**
	(3.86)	(6.08)	(6.63)	(6.11)	(6.25)
	[0.04]	[0.21]	[0.36]	[0.38]	[0.34]
	Sub-sample 1981Q1-2013Q3, recursive				
cay^{FC}	0.17	1.00	2.48	3.96	6.39
	(0.48)	(0.83)	(1.04)	(1.18)	(1.82)
	[-0.01]	[0.00]	[0.03]	[0.06]	[0.11]
cay^{FCrec}	0.30	1.67	**4.04**	**6.16**	**8.10**
	(0.97)	(1.65)	(2.29)	(2.79)	(4.17)
	[0.00]	[0.04]	[0.16]	[0.27]	[0.41]
cay^{MSrec}	0.41	2.13	**6.01**	**8.65**	**10.33**
	(1.10)	(1.92)	(2.73)	(3.51)	(5.17)
	[0.00]	[0.04]	[0.21]	[0.31]	[0.37]

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

Long Horizon Forecasting Regressions: Stock Returns - PCE *cay*

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}	**0.78**	**3.34**	**6.51**	**8.41**	**9.45**
	(3.06)	(3.78)	(5.32)	(6.62)	(7.99)
	[0.03]	[0.14]	[0.30]	[0.39]	[0.44]
cay^{MSfilt}	**0.95**	**4.12**	**6.55**	**7.61**	**8.53**
	(3.18)	(3.36)	(3.46)	(3.07)	(3.31)
	[0.03]	[0.13]	[0.19]	[0.19]	[0.21]
cay^{MS}	**0.96**	**4.15**	**6.49**	**7.80**	**8.60**
	(3.15)	(3.39)	(3.40)	(3.17)	(3.42)
	[0.03]	[0.13]	[0.18]	[0.20]	[0.21]
	Sub-sample 1981Q1-2013Q3, recursive				
cay^{FC}	0.73	**3.28**	**6.99**	**9.51**	**10.01**
	(1.88)	(2.44)	(3.03)	(3.66)	(3.86)
	[0.01]	[0.08]	[0.24]	[0.31]	[0.29]
cay^{FCrec}	0.18	1.08	**2.95**	**4.39**	**5.28**
	(0.96)	(1.35)	(1.98)	(2.36)	(2.89)
	[0.00]	[0.02]	[0.12]	[0.20]	[0.25]
cay^{MSrec}	**1.41**	**5.75**	**7.21**	**7.72**	**6.72**
	(2.80)	(3.61)	(3.96)	(3.68)	(2.95)
	[0.05]	[0.23]	[0.21]	[0.16]	[0.10]

Table 6: This tables reports the results from regressions of of h-period-ahead CRSP-VW returns in excess of a 3-month Treasury-bill rate, $r_{f,t}$, on the variable listed in the first column. cay^{FC} is the fixed-coefficient consumption-wealth ratio; cay^{MSfilt} denotes the Markov-switching version of *cay* using filtered probabilities and cay^{MS} denotes the benchmark Markov-switching *cay* using smoothed probabilities. The bottom panel reports results from regressions using recursively estimated versions of *cay*, in which all parameters are estimated using data up to time t rather than using the full sample. The models are first estimated on data from 1952Q1-1970Q1. We then recursively add observations and reestimate the *cay* variables over expanding sub-samples using data only up to the end of that subsample, continuing in this way until the end of the sample, 2013:Q3. Results are reported for the subsample since 1980. cay^{FCrec} denotes the fixed coefficient *cay* estimated recursively, while cay^{MSrec} denotes the Markov-switching *cay* estimated recursively using smoothed probabilities. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses), and adjusted R^2 statistics in square brackets. Significant coefficients based on a t-test at the 5% significance level are highlighted in bold face. The full sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Out-Of-Sample Forecasts

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

$z_t =$	Horizon h (in quarters) 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
	NDS consumption				
cay^{FC}	0.71	2.90	4.67	6.74	7.36
cay^{MSfilt}	0.70	2.47	2.64	3.01	3.72
cay^{MS}	0.70	2.35	2.53	2.92	3.68
cay^{FCrec}	0.72	2.87	4.38	5.72	6.61
cay^{MSrec}	0.71	2.86	4.49	5.75	6.14
	PCE consumption				
cay^{FC}	0.71	2.63	3.56	4.81	5.82
cay^{MSfilt}	0.68	2.09	3.34	5.54	7.89
cay^{MS}	0.68	2.09	3.34	5.47	7.80
cay^{FCrec}	0.66	2.16	4.43	6.99	9.07
cay^{MSrec}	0.72	2.84	4.29	5.51	6.38

Table 7: This tables reports the mean-squared forecast errors from out-of-sample h-period-ahead forecasts of CRSP-VW returns in excess of a 3-month Treasury-bill rate using 60-quarter rolling subsamples. The single predictor variable in each regression is listed in the first column. The forecasting regression is first estimated on data from 1952Q1-1980Q1, and forecasts are made over the next h periods. We then repeat this forecasting regression using data from the next 60 quarters of the sample, continuing in this way until the end of the sample, 2013:Q3. Mean-square-errors are reported for the subsample since 1980. cay^{FC} is the fixed-coefficient consumption-wealth ratio, cay^{MSfilt} and cay^{MS} are the Markov-switching cay variables using filtered and smoothed probabilities, respectively, cay^{FCrec} is the recursively estimated cay with fixed coefficients, and cay^{MSrec} is the recursively estimated Markov-switching cay. The recursive estimates use data only up to time t. The full sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

Summary statistics for the Real Interest Rate and GDP growth

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

Table 8: This table reports the mean and standard deviation for the real interest rate and GDP growth based on the VAR estimates conditional on staying in each regime.

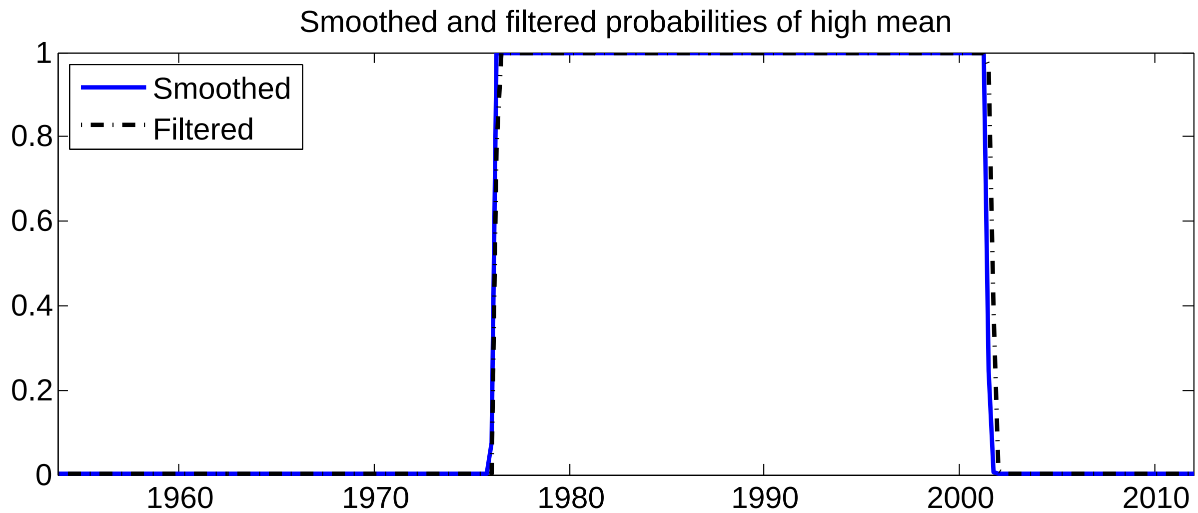

Figure 1: Smoothed probability of high mean regime for the Markov-switching cointegrating relation. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

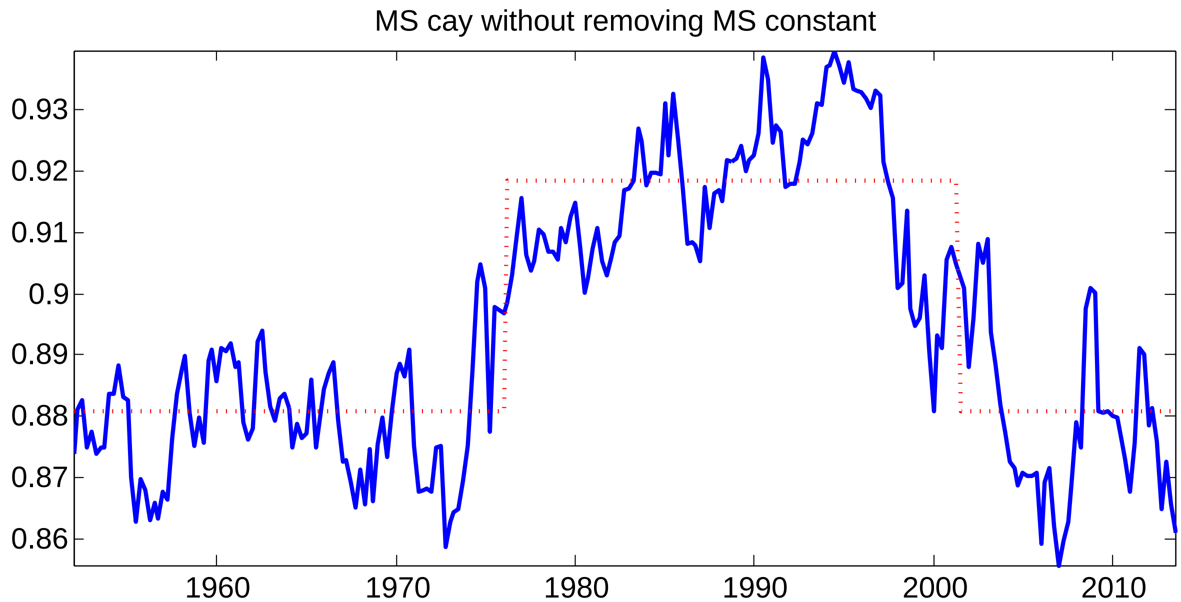

Figure 2: The Markov-switching estimated cay^{MS} is plotted with*out* removing the constant. The red dashed lines are the values of α_1 and α_2, which correspond to the most likely mean values in each regime. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

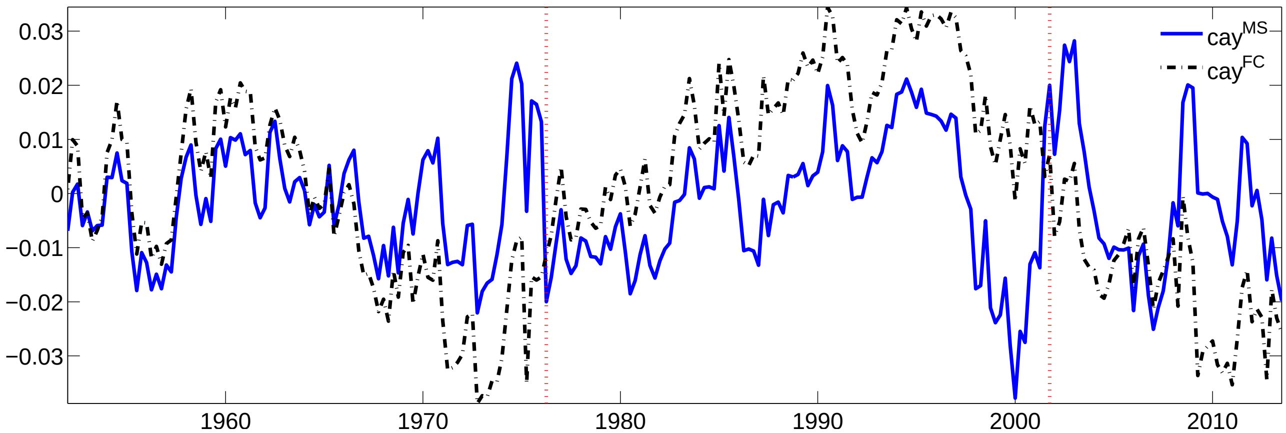

Figure 3: Markov-switching and fixed coefficients *cay*. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

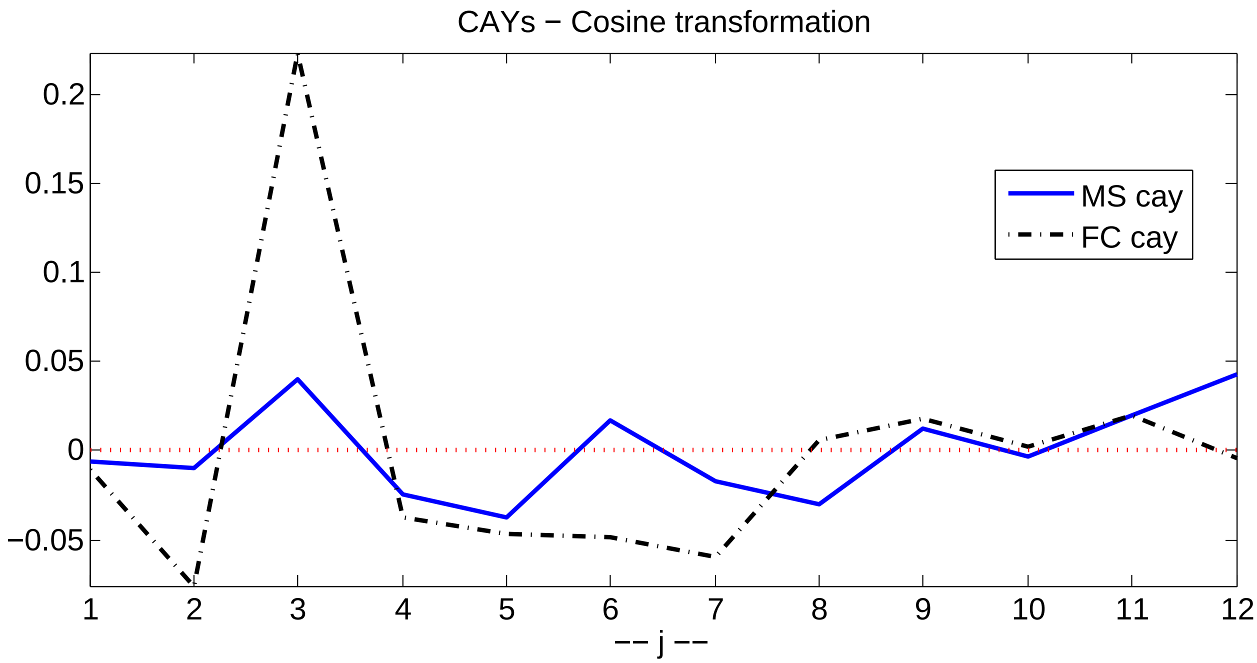

Figure 4: Low frequency averages of *cay*. The figure plots the set of averages $\{f_j\}_{j=1}^{k}$, which capture the variability in *cay* for periods greater than $2T/k$, where T is the sample size. Thus, with T $=$ 247 quarters, the $k = 12$ points plotted summarize the variability in *cay* for periods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

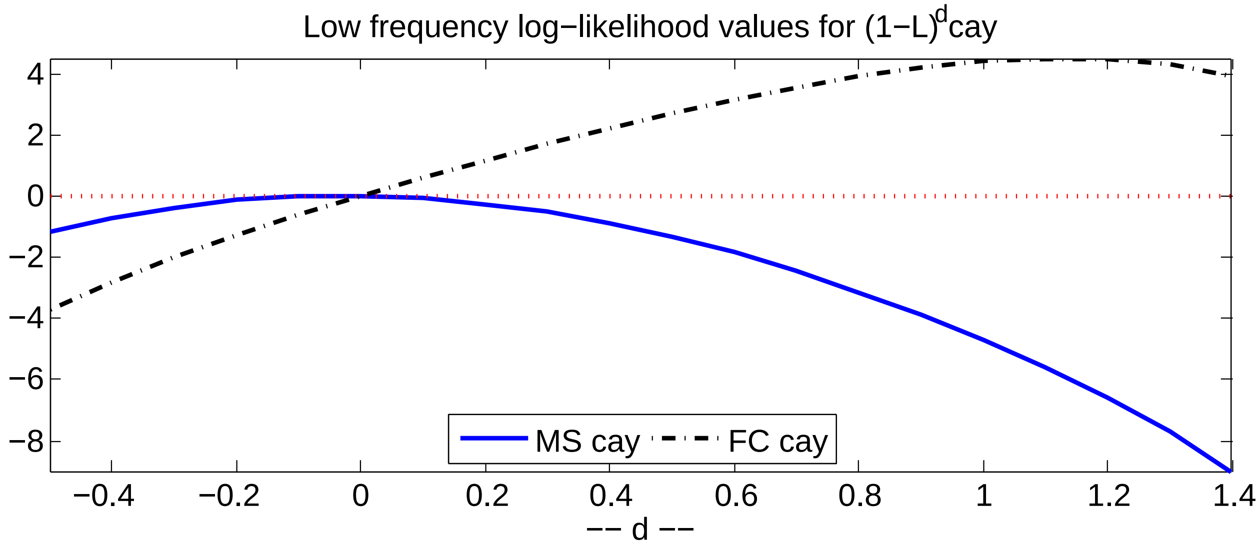

Figure 5: Low frequency log likelihood values for $(1-L)^d cay_t$. The sample is quarterly and spans the period 1952:Q1 to 2013:Q3.

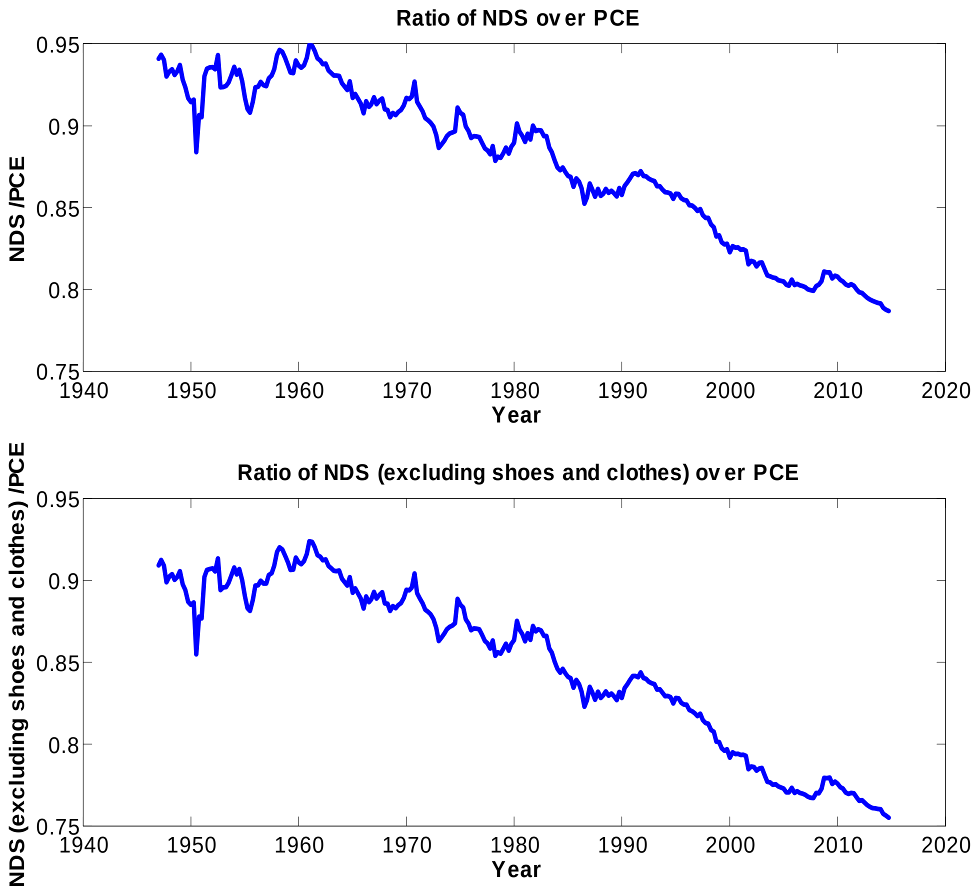

Figure 6: **NDS Ratio.** Real total PCE (including durable, service and nondurable) is obtained directly from BEA, measured 2009 chain-weighted dollars with 1999 as base year. NDS expenditures are chain-weighted together appropriately using the same deflator. The sample spans the period 1952:Q1 to 2014:Q3.

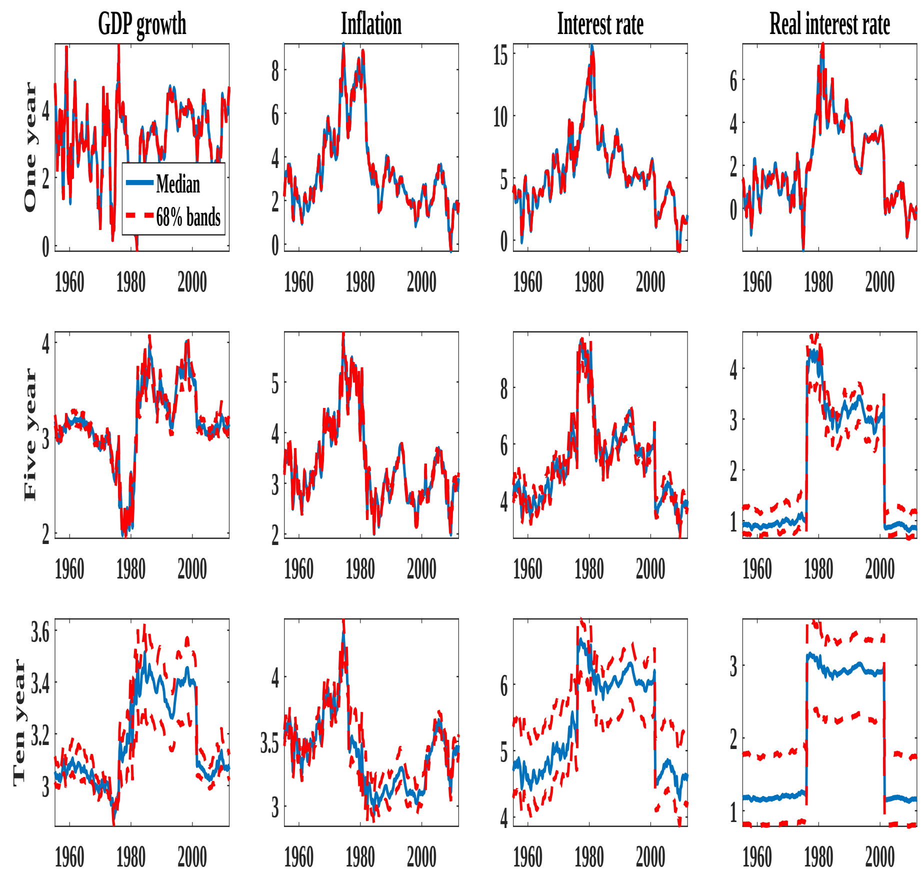

Figure 7: **Projections from MS-VAR.** The figure reports the conditional expectations based on the MS-VAR at different horizons taking into account the possibility of regime changes.

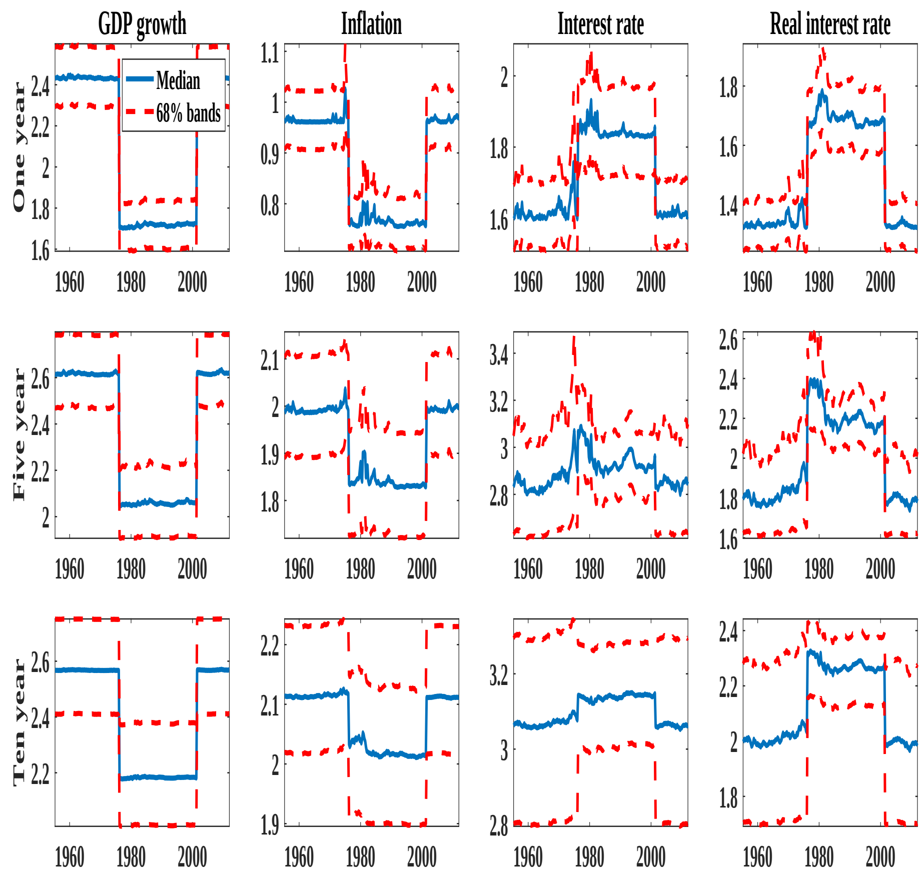

Figure 8: **Uncertainty based on MS-VAR.** The figure reports the conditional standard deviations at different horizons based on the MS-VAR taking into account the possibility of regime changes.

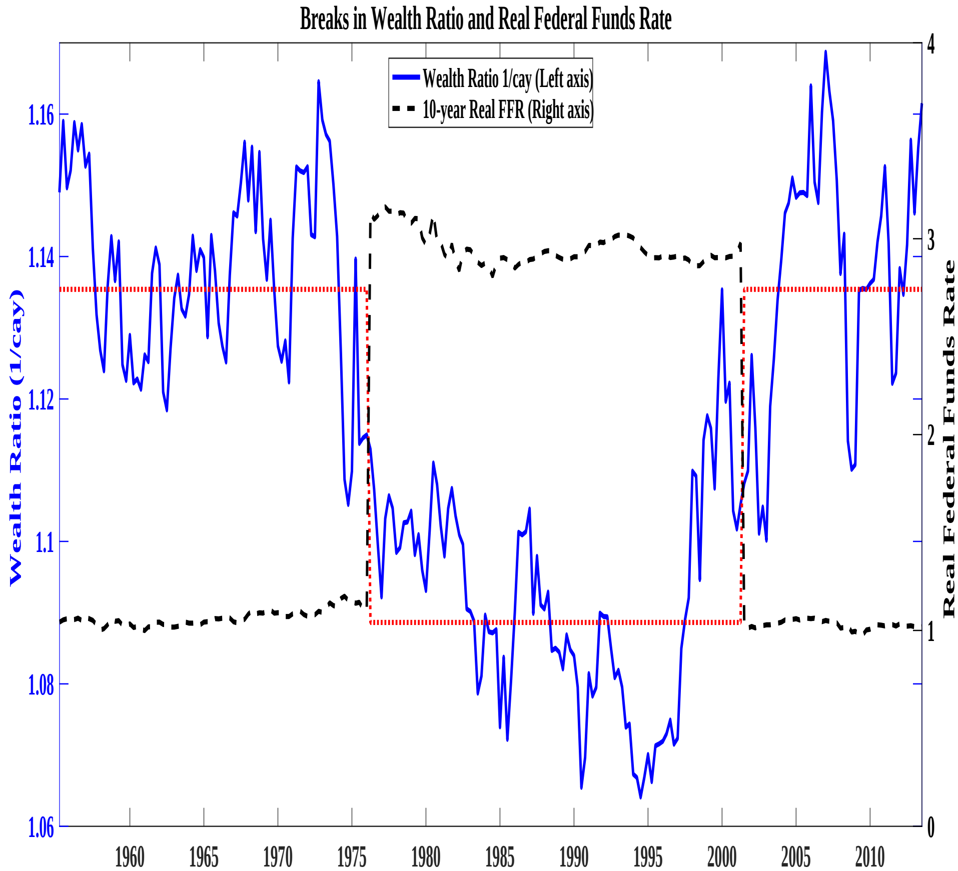

Figure 9: Wealth Ratio and federal funds rate. The wealth ratio (solid blue line, left axis) is plotted together with the ten-year-ahead real federal funds rate (black dashed line, right axis). The wealth ratio is obtained as the inverse of cay^{MS} without removing the Markov-switching constant. The red dashed line represents the inverse of the regime-probability weighted average of the constants α_1 and α_2. The ten-year-ahead real federal funds rate is computed as the ten-year-ahead expected value of the real federal funds fate as implied by the Markov-switching VAR. The sample is quarterly and spans the period 1955:Q4 to 2013:Q3. With respect to the estimates for cay^{MS} the sample is adjusted to take into account data availability for the Federal Funds rate.