THE FEDERAL RESERVE BANK
of ST. LOUIS

WORKING PAPER SERIES

Non-Markovian Regime Switching with Endogenous States and Time-Varying State Strengths

Siddhartha Chib
and
Michael Dueker

_____

Non-Markovian Regime Switching with Endogenous States and Time-Varying State Strengths

October 2004

Siddhartha Chib
Olin School of Business
Washington University
chib@olin.wustl.edu

Michael Dueker*
Federal Reserve Bank of St. Louis
P.O. Box 442, St. Louis, MO 63166
mdueker@stls.frb.org; fax (314) 444-8731

I

Non-Markovian Regime Switching with Endogenous States and Time-Varying State Strengths

Abstract

This article presents a non-Markovian regime switching model in which the regime states depend on the sign of an autoregressive latent variable. The magnitude of the latent variable indexes the 'strength' of the state or how deeply the system is embedded in the current regime. In this model, regimes have dynamics, not only persistence, so that one regime can gradually give way to another. In this framework, it is natural to allow the autoregressive latent variable to be endogenous so that regimes are determined jointly with the observed data. We apply the model to GDP growth, as in Hamilton (1989), Albert and Chib (1993) and Filardo and Gordon (1998) to illustrate the relation of the regimes to NBER-dated recessions and the time-varying expected durations of regimes. The article makes use of the Metropolis-Hastings algorithm to make multi-move draws of the latent regime strength variable, where the extended Kalman filter provides a valid proposal density for the latent variable.

JEL classifications: **F42, C25, C22**

Key words: **Regime switching, Markov Chain Monte Carlo, nonlinear state space**

II

Introduction

Autoregressive models are popular in economics because many economic variables appear to respond more to their own past values than they do to a distributed lag of any other variable. The same is likely true of regimes. If conditions gradually become ripe for a regime change, it might not be possible to find an exogenous covariate whose evolution matches this ripening process. For example, when modeling the Volcker monetary policy regime change in 1979, a regime modeler might claim that the occurrence of high inflation engendered a shift in probability toward a new regime. If this were true, then a regime-switching model could include past inflation as an explanatory variable in Filardo's (1994) time-varying transition probability Markow switching model. However, the history of monthly or quarterly inflation rates (or almost any other extrinsic variable) does not suggest a uniquely high monthly inflation rate that served as a trigger for a change of monetary policy regime in 1979 [see Sims and Zha (2004) for a discussion of monetary policy regimes]. Instead, it is likely that pressure for a regime change built gradually across time Similarly, if the regime studied is the recession/expansion state of the business cycle, then a well-known problem is how to identify a variable or set of variables that heralds a shift from an expansion phase to a recession phase. In both of these cases, autoregressive dynamics might prove more useful than a distributed lag of any exogenous covariate in modeling a gradual shift in regime probabilties.

A related issue is the extent to which regimes are determined separately from the observable data. A negative shock that moves the business cycle phase toward the expansion state from the recession state is possibly associated with a negative shock to observed GDP growth. These two shocks do not have to be postively correlated, and they might even be negatively correlated, but regime modelers should be hesitant to assume that the regime is exogenous and uncorrelated with the innovations to the data affected by the

regime. With our latent variable approach, it is quite simple and straightforward to allow for endogenous regimes that are correlated with the observable data.

This article introduces a new non-Markovian regime switching model in which the regime states depend on the sign of an autoregressive latent variable. The magnitude of the latent variable indexes the 'strength' of the state or how deeply the system is embedded in the current regime. This non-Markovian regime switching automatically implies time-varying state transition probabilities. With autoregressive dynamics governing the transition probabilities, we can readily demonstrate how the expected duration of the current regime can vary across time. In this way, the regimes themselves have autoregressive dynamics, so that pressure for a regime change can build gradually across time. In essence, our model of hidden regimes is the counterpart to the dynamic probit approach to observed regimes, as discussed below. This model is readily contrasted with the typical two-state Markov switching model if we write the transition probabilities of a two-state Markov process as a function of a normally-distributed latent variable, S^*, that governs the binary regime indicator S:

$$
\begin{aligned}
S_t^* &= \lambda + \theta S_{t-1} + e_t \\
e_t &\sim N(0, 1) \\
S_t &= 0 \quad \Longleftrightarrow \quad S_t^* < 0
\end{aligned}
\tag{1}
$$

The constant transition probabilities for this Markov process are therefore parameterized as

$$
\begin{aligned}
P(S_t = 0 \mid S_{t-1} = 0) &= \Phi(-\lambda) \\
P(S_t = 1 \mid S_{t-1} = 1) &= 1 - \Phi(-\lambda - \theta),
\end{aligned}
\tag{2}
$$

where $\Phi(.)$ is the cumulative standard normal density function. With constant transtition probabilities, the Markov switching model implies a constant expected duration of the current regime. Note that, when the regimes are observed, models of the form of eq. (1) are sometimes inappropriately called dynamic probits [DeJong and Woutersen (2004); Horowitz (1992)], because the same nomenclature is also used to describe the dynamic probit model of Eichengreen, Watson and Grossman (1985), where the lagged latent variable is on the right-hand side. Putting the lagged state on the right-hand side adds persistence, not dynamics, to the regimes. To see this, note that eq. (1) is equivalent to:

$$
\begin{aligned}
S_t^* &= \lambda + e_t \\
S_t &= 0 \text{ if } S_t^* < 0, S_{t-1} = 0 \\
&\quad \text{or } S_t^* < -\theta, S_{t-1} = 1
\end{aligned}
\tag{3}
$$

The form of equation (3) makes clear that the regime strength, S^*, has no dynamics. For this reason, the transition probabilities in Markov switching models are called persistence parameters because they do not connote regime dynamics.

The model we propose with autoregressive state strengths takes the form

$$
\begin{aligned}
S_t^* &= \lambda + \theta S_{t-1}^* + e_t \\
e_t &\sim N(0, 1) \\
S_t &= 0 \iff S_t^* < 0
\end{aligned}
\tag{4}
$$

This model is the hidden regime counterpart to the dynamic probit model of Eichengreen, Watson and Grossman (1985) because the latent variable is autoregressive, implying true regime dynamics. This autoregressive latent variable generates a non-Markovian regime process because the probability of the state this period depends not only on the state

last period but a continuous measure of the strength of the state last period. For these non-Markovian regimes, the time-varying state transition probabilities are

$$
\begin{aligned}
P(S_t = 0 \mid S_{t-1}^*) &= \Phi(-\lambda - \theta S_{t-1}^*) \\
P(S_t = 1 \mid S_{t-1}^*) &= 1 - \Phi(-\lambda - \theta S_{t-1}^*).
\end{aligned}
\tag{5}
$$

In the model with autoregressive state strengths, the probability of a regime change would rise if the latent index of regime strength, S^*, approached zero.

The Markov switching model time-varying transition probabilities introduced by Filardo (1994) would add lagged covariates, Z, such that

$$
S_t^* = \lambda + \theta S_{t-1} + \kappa Z_{t-1} + e_t
\tag{6}
$$

$$
e_t \sim N(0,1)
$$

$$
S_t = 0 \iff S_t^* < 0.
$$

Our model with autoregressive state strengths suggests that one strong candidate to be included in Z_{t-1} is S_{t-1}^*, whereupon S_{t-1} becomes unnecessary.

As mentioned above, one useful feature of the non-Markovian reqime switching model of equation (4), unlike the Markov switching model of equation (1), is that the expected duration of a regime is time-varying. Filardo and Gordon (1998) add time-varying expected regime durations by way of time-varying transition probabilities. In general, this covariate approach to time-varying expected durations requires an auxiliary model to predict the future evolution of the Z covariates. Lam (2004) uses the regime durations as Z_{t-1} in eq. (6) to make the transition probabilities vary across time, where the unobserved regimes are counted relative to a probability threshold. Our autoregressive model of regime strengths, in contrast, implies time-varying expected regime durations even in the absence of covariates in the regime equation.

II. MCMC estimation of non-Markovian regime switching

We estimate the model with a latent autoregressive variable via Markov Chain Monte Carlo methods. MCMC methods for estimating the hidden Markov switching model of Hamilton (1989) were put forth in Albert and Chib (1993), who showed that once one augments the data with draws of the latent regime states, then the conditional distributions of the other model parameters are straightforward regression coefficient priors and posteriors.

The specific model that we apply to GDP growth, denoted y, is

$$
\begin{aligned}
y_t &= \alpha_1 + (\alpha_0 - \alpha_1)I(S_t^* < 0) + \phi y_{t-1} + u_t \\
S_t^* &= \lambda + \theta_1 S_{t-1}^* + \theta_2 y_{t-1} + e_t
\end{aligned}
\tag{7}
$$

where $I(.)$ is the indicator function. Note that in this specification only the nonlinearity of the indicator function (and, more specifically, the cumulative density function, which is the forecast of the indicator function) allows one to separately identify ϕ and θ_2, for example. This identification is potentially sensitive to the distributional assumption one uses for the cumulative density function, as discussed by Heckman and Macurdy (1986). Therefore, we compare results where the identification relies on a distributional assumption, as in eq. (7) above, and a specification where a variable other than y_{t-1} appears on the right side of the latent state equation. In the latter case, an independent source of variation in the latent variable S^* is ensured and nonlinearlity is no longer the sole source of identification.

The covariance matrix of the error terms is allowed to be general such that

$$
\text{Cov}\left(\begin{array}{c} u_t \\ e_t \end{array} \right) = \Sigma = \left(\begin{array}{cc} \sigma_u^2 & \rho \\ \rho & 1 \end{array} \right).
\tag{8}
$$

The parameter groupings for MCMC estimation of the model are

$$
\varrho_1 = (\alpha_0, \alpha_1, \phi)
$$

$$\varrho_2 = (\lambda, \theta_1, \theta_2)$$

$$\varrho_3 = \Sigma$$

$$\varrho_4 = \{S_t^*\}, t = 1, ..., T \tag{9}$$

Because the model of equation (6) is easily cast in state-space form as

$$y_t = \alpha_1 + (\alpha_0 - \alpha_1)I(S_t^* < 0) + \rho e_t + v_t$$

$$\begin{pmatrix} e_{t+1} \\ S_t^* \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & \theta_1 \end{pmatrix} \begin{pmatrix} e_t \\ S_{t-1}^* \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \lambda & \theta_2 \end{pmatrix} \begin{pmatrix} 1 \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{t+1} \\ 0 \end{pmatrix}, \tag{10}$$

a natural estimation approach to consider is to use the Kalman filter to integrate out the unobserved latent state strength variable S^*. Instead of integrating out S^*, however, we choose to sample it for two reasons: First, this state-space model is nonlinear and the extended Kalman filter applied to nonlinear state-space models is inexact and, therefore, we subject the draw of the latent variable to a Metropolis-Hastings step; second, the conditional distributions of the model parameters in ϱ_1 and ϱ_2 are easily derived from simple regressions conditional on value of S^*. Thus, we find it convenient to make use of the data augmentation capabilities of Markov Chain Monte Carlo methods.

To use regression techniques to derive a conditional mean and variance for ϱ_1, it is necessary to control for the endogeneity of S^*. Fortunately, the data augmentation makes this relatively simple. [1] Conditional on $(\varrho_2, \varrho_3, \varrho_4)$, we can write

$$u_t = \rho e_t + v_t, \tag{11}$$

where v_t is uncorrelated with e_t, and re-write equation (7) as

$$y_t - \rho e_t = \alpha_0 I(S_t^* < 0) + \alpha_1 I(S_t^* \geq 0) + \phi y_{t-1} + v_t \tag{12}$$

---

[1]Kim, Piger and Startz (2004) discuss maximum-likelihood estimation of a Markov switching model with endogenous regimes that does not involve data augmentation.

In this form, we have a regression equation in which the error term is uncorrelated with the regressor $I(S_t^* < 0)$. The conditional distribution of the coefficients in ϱ_1 is Normal with the mean and variance implied by the Bayesian regression, given a prior. The priors used for the Bayesian regressions are discussed in section III.

Similarly, the regime strength equation must take account of the non-zero $E(e_t \mid u_t)$ by writing

$$e_t = \rho/\sigma_u^2 u_t + \nu_t.$$

The regression equation thus becomes

$$S_t^* - \rho/\sigma_u^2 u_t = \lambda + \theta_1 S_{t-1}^* + \theta_2 y_{t-1} + \nu_t \tag{13}$$

Conditional on $\varrho_1, \varrho_2, \varrho_4$, the residual series $\{u_t\}$ and $\{e_t\}$ are calculated and the approach from Chib, Greenberg and Jeliazkov (2003) is used to sample the covariance matrix Σ using inverted Wishart distributions, subject to the restriction that $\Sigma_{2,2} = 1$. A detailed discussion of sampling the autoregressive latent variable follows.

Sampling the latent variable

To reduce the degree of autocorrelation of the sampled values across MCMC iterations and to speed convergence of the sampler to the posterior distribution, multi-state sampling is preferable to single-state sampling of the latent variable. In single-state sampling, the conditional distribution of the latent variable this iteration would depend on values drawn for the previous iteration (with iteration number denoted as a superscript):

$$f(S_t^{*(i+1)} \mid \{S_j^{*(i+1)}\}_{j<t}, \{S_k^{*(i)}\}_{k>t}, \{y_t\}). \tag{14}$$

In our application, single-move sampling appeared not to converge, even after we discarded more than 100,000 burn-in iterations. The single-move posterior means differed

substantially across estimation runs, whereas the multi-move sampler reproduces nearly identical results across numerous estimation runs of 40,000 iterations, each with 10,000 discarded. As suggested by Carter and Kohn (1994), Fruhwirth-Schnatter (1994) and De Jong and Shephard (1995), multi-state sampling can be carried out based on the identity

$$f(\{S_t^*\} \mid \{y_t\}) = f(S_T^* \mid \{y_t\}) \prod_{t=1}^{T-1} f(S_t^* \mid S_{t+1}^*, \{y_j\}_{j=1,...,t}), \tag{15}$$

using the Kalman filter to calculate the conditional distributions on the right side of eq. (14). One key feature of our approach, however, is that we use the extended Kalman filter only to produce a proposal density for the latent state strength index, S^*, and not to claim that the filter gives an exact conditional distribution [see Welch and Bishop (2002) for a useful summmary of the extended Kalman filter].

If we start with a canonical linear state-space model with observation variables y and state variables X,

$$
\begin{aligned}
y &= HX_t + v_t \\
X_{t+1} &= FX_t + DZ_t + w_{t+1},
\end{aligned}
\tag{16}
$$

then the well-known Kalman filtering equations are

$$
\begin{aligned}
X_{t+1|t} &= FX_{t|t} + DZ_t \\
X_{t+1|t+1} &= X_{t+1|t} + K_{t+1}[y_{t+1} - HX_{t+1|t}] \\
P_{t+1|t} &= FP_{t|t}F' + Q \\
P_{t+1|t+1} &= P_{t+1|t} - K_{t+1}HP_{t+1|t} \\
K_{t+1} &= P_{t+1|t}H'(HP_{t+1|t}H' + R)^{-1}
\end{aligned}
\tag{17}
$$

As shown above, the non-Markovian regime switching model has the nonlinear state-space form of eq. (10). The extended Kalman filter is based on approximating the

8

nonlinear functions in the state-space model. In the case of eq. (10), the nonlinear function is the indicator function which we approximate as

$$I(S^* < 0) \approx P(S^* < 0 \mid I_{t-1}) + H_t(e_t, S_{t-1}^* - S_{t-1|t-1}^*)', \tag{18}$$

which is superior to the Taylor series approximation

$$I(S^* < 0) \approx I(S_{t|t-1}^* < 0) + H_t(e_t, S_{t-1}^* - S_{t-1|t-1}^*)'. \tag{19}$$

In most applications of the extended Kalman filter, the Taylor series approximation is used because typically the nonlinear function is one where we know how to take expectations of its arguments, but we do not now how to take expectations of the function itself. With the indicator function, however, we can take the expectation of the function directly as the cumulative density function, as in eq. (18), and avoid the Jensen's inequality problem that plagues the Taylor series approximation. It also puts the $P(S_t^* < 0 \mid I_{t-1})$, which is the most natural forecast of $I(S_t^* < 0)$, into the Kalman gain equation.

The extended Kalman filter calls for replacement of the indicator in eq. (7) with the approximating eq. (18) and replaces the vector H from eq. (17) with the the Jacobian, H_t, of the approximating eq. (18). In this case, the 1×2 Jacobian vector, H_t, includes a finite-difference approximation to differentiating the indicator function, where $X_t = (\epsilon_t, S_{t-1}^*)'$:

$$H_t = \frac{E\Delta I(S^*(X_t) < 0)}{E[\xi_t \mid S_t^* \gtrless 0, S_{t-1}^* \lessgtr 0]} \frac{\partial \xi_t}{\partial X_t} - \frac{\partial P(S^*(X_t) < 0 \mid I_{t-1})}{\partial \xi_t} \frac{\partial \xi_t}{\partial X_t} < 0, \tag{20}$$

where

$$\xi_t = \frac{S_t^* - \lambda - \theta_1 S_{t-1|t-1}^* - \theta_2 y_{t-1}}{(P_{t|t-1}^{(1,1)} - 2\theta_1 P_{t|t-1}^{(1,2)} + \theta_1^2 P_{t|t-1}^{(2,2)})^{0.5}} \tag{21}$$

and superscript (2,2) indicates the element of the matrix. The negative sign on H_t reflects the expected decrease in the indicator function from a shock to the unobserved components of the latent variable S^*. Because ξ_t is a standard normal,

$$E[\xi_t \mid S_t^* > 0] = \phi(\xi_t \mid S_t^* = 0)/[1 - \Phi(\xi_t \mid S_t^* = 0)]$$

9

$$
\begin{aligned}
E[\xi_t \mid S_t^* < 0] &= -\phi(\xi_t \mid S_t^* = 0)/\Phi(\xi_t \mid S_t^* = 0) \\
\frac{E\Delta I(S^*(X_t) < 0)}{E[\xi_t \mid S_t^* \gtrless 0, S_{t-1}^* \lessgtr 0]} &= -1[1 - \Phi(\xi_t \mid S_t^* = 0)]I(S_{t-1|t-1}^* < 0)/E[\xi_t \mid S_t^* > 0] \\
&\quad +\Phi(\xi_t \mid S_t^* = 0)I(S_{t-1|t-1}^* > 0)/E[\xi_t \mid S_t^* < 0] \\
\frac{\partial \xi_t}{\partial X_t} &= \begin{pmatrix} 1 \\ \theta_1 \end{pmatrix}'/(P_{t|t-1}^{(1,1)} - 2\theta_1 P_{t|t-1}^{(1,2)} + \theta_1^2 P_{t|t-1}^{(2,2)})^{0.5} \quad (22)
\end{aligned}
$$

where $\phi(.)$ is the standard normal density function and $\Phi(.)$ is the cumulative standard normal density. The ratio of finite differences in H_t represents the probability that a shock to the latent variable will cause a change in the indicator function times the sign of the change in the indicator function divided by the expected value of the shock conditional on it being large enough to induce a regime change.

Thus, the extended Kalman filtering equations are altered from the canonical form of eq. (17) to

$$
\begin{aligned}
X_{t+1|t} &= FX_{t|t} \\
X_{t+1|t+1} &= X_{t+1|t} + K_{t+1}[y_{t+1} - \phi y_t - \alpha_1 - (\alpha_0 - \alpha_1)P(S^*(X_{t+1}) < 0 \mid I_t)] \\
P_{t+1|t} &= FP_{t|t}F' + Q \\
P_{t+1|t+1} &= P_{t+1|t} - K_{t+1}H_{t+1}P_{t+1|t} \\
K_{t+1} &= P_{t+1|t}H_{t+1}'(H_{t+1}P_{t+1|t}H_{t+1}' + R)^{-1} \quad (23)
\end{aligned}
$$

where

$$
P(S^*(X_{t+1}) < 0 \mid I_t) = \Phi(\xi_{t+1} \mid S_{t+1}^* = 0). \quad (24)
$$

We also need to apply one smoothing step following the sampling of X_{t+1}:

$$
\begin{aligned}
X_{t|t+1} &= X_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}(X_{t+1} - X_{t+1|t}) \\
P_{t|t+1} &= P_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}(P_{t+1|t+1} - P_{t+1|t})(P_{t|t}F'P_{t+1|t}^{-1})' \quad (25)
\end{aligned}
$$

In this way, the latent variable is sampled in reverse order, starting with $X_T = (\epsilon_T, S^*_{t-1})'$.

It is important to note that this distribution for the latent variable is only considered a proposal density and these draws of the latent variable are a draw from the proposal density. In a Metropolis-Hastings step, the proposal density does not have to represent the exact conditional distribution of the parameters, which the extended Kalman filter does not provide. Instead, the proposal density simply needs to provide a useful approximation to the posterior density of the parameters in question. The posterior density of S^* can be evaluated directly via Bayes' Law and does not involve Kalman filter recursions at all.

The latent variable vector, $\{S^*_t\}, t = 1, ..., T$, is updated according to the following AR-MH algorithm:

1. Draw a proposed value of the vector $\{S^*_t\}$, using as a proposal density the distribution implied by the extended Kalman filtering algorithm outlined above. Denote this proposal density as $q(.)$.

2. Given an uninformative prior for the latent variable, the posterior density of $\{S^*_t\}$ depends only on the density of the data conditional on the value of $\{S^*\}$. We need to calculate the densities of the data conditional on the proposed value of $\{S^*\}$ and conditional on last iteration's value of $\{S^*\}$, denoted $h(S^{*p})$ and $h(S^{*c})$, respectively.

3. The acceptance probability for S^{*p}, as opposed to staying at S^{*c}, is

$$min\{\frac{h(S^{*p})q(S^{*c})}{h(S^{*c})q(S^{*p})}, 1\}.$$

The efficiency of the Metropolis-Hastings sampler depends greatly on the acceptance rate of the proposed draws. Our method of drawing the latent variable vector $\{S_t\}$ resulted in an acceptance rate of approximately 75 percent (54 percent in the model specification that uses the leading indicators as an explanatory variable in the latent

11

regime strength equation) in our application to GDP growth. This high rate is a sign that our modification to the extended Kalman filter is leading to highly useful inferences of the latent state strength measure. We discuss the ability of our algorithm to track the latent variable in Mote Carlo simulations below.

III. A Monte Carlo investigation of the sampling procedure

To investigate how well the extended Kalman filter uncovers the parameters of the data-generating process for the latent variable, we performed a Monte Carlo simulation. We generated 1000 samples of artificial data, each with 200 observations, based on eq. (7). In this form, with y_{t-1} as a covariate in the latent state strength equation, the system is self-contained and can be simulated without additional assumptions. The true parameter values and priors were set close to those reported below for the application to GDP growth. The regression priors are discussed in the next section with the GDP results.

For each sample, we ran the MCMC estimation procedure for 25000 iterations and we saved the posterior means of the parameter draws from the last 20000 iterations. For each estimation, we saved the 5, 50 and 95 percent quantiles for each parameter and then calculated the average of these quantiles across the 1000 estimations. Table 1 shows the results from this Monte Carlo investigation. In all cases, the true parameter value lies comfortably within the estimated 90 percent interval. In addition, the 50 percent quantile can serve as a useful point estimate of the parameter value. From this monte carlo exercise, we also saved quantiles of approximately every 15th value of the latent variable, S^*. More specifically, we saved the 5, 50 and 95 percent quantiles of the difference between the MCMC inferred values and true values of the latent variable. Figure 1 depicts the 90

percent intervals, which are fairly wide for a given observation, but it also shows that there is no significant bias in the sampling algorithm.

Table 1: Monte Carlo Simulation of MCMC Sampling Algorithm			
	Inferred Quantiles	**True value**	**Prior**
	Observation equation		
α_0	0.131	0.10	0.10
	(0.036,0.226)		
α_1	0.873	0.90	0.90
	(0.780,0.966)		
ϕ	0.329	0.30	None
	(0.237,0.420)		
	Latent regime equation		
λ	0.341	0.30	None
	(0.040,0.752)		
θ_1	0.549	0.60	0.70
	(0.400,0.698)		
θ_2	-0.002	0.05	None
	(-0.178,0.183)		
	Covariance matrix		
σ_1^2	0.888	0.80	None
	(0.706,1.11)		
ρ	0.225	0.30	None
	(0.043,0.467)		
Estimated 50 percent quantiles with 5 and 95 percent quantiles in parentheses			

IV. Application to business cycle phases

In applying this regime switching model to GDP growth, we found that an informative prior is necessary to slow down the fluctuations in the latent variable. With an uninformative prior, the inferred latent S^* series closely mimics the data y with a different mean and variance. The estimated values of growth states, α_0 and α_1, are also closer together than one would associate with two distinct business cycle regimes in the absence of an

informative prior. In a Bayesian regression, as shown in Chib and Greenberg (1996), the coefficients are normally distributed such that:

$$\beta \sim N(\hat{\beta}, B_n^{-1}) \tag{26}$$
$$B_n = B_0 + X'X/\sigma^2$$
$$\hat{\beta} = B_n^{-1}(B_0\beta_0 + X'y/\sigma^2),$$

where X is set of regressors, y is the regressand, σ^2 is the variance and, most importantly, B_0 is a diagonal matrix that determines the strength of the prior placed on the set of coefficient values β_0. For the GDP growth regression of eq. (7), where the coefficients are $(\alpha_0, \alpha_1, \phi)$, the diagonal elements of B_0 were set to (300,300,0) and β_0 was set to (0.10,0,90,0.30), so that no prior was placed on the lagged dependent variable. For the latent state from eq. (7), where the coefficients are $(\lambda, \theta_1, \theta_2)$, the diagonal elements of B_0 were set to (0,100,0) and β_0 was set to (0.30,0.80,0.10), so that the prior only served to lift the autoregressive coefficient, θ_1. Experimentation showed that these priors were strong enough to prevent the regime from changing in more than one-third of the observations; when the regime changes more often than this, the model is trying to fit high-frequency fluctuations between two expansionary growth states, as opposed to lower-frequency business cycle fluctuations.

One obvious question is why the Markov switching regimes of Hamilton (1989) do not require any prior restrictions in order to match business cycle fluctuations, whereas the present non-Markovian regime switching model does. Consider first the Markov switching model. Suppose that it tried to fit high-frequency fluctuations between two expansionary growth states of 2.5 and 4 percent annualized growth. With fixed transition probabilities the model would need to have states that were not very persistent to have relatively frequent transitions. As a consequence though, the one-step-ahead forecast of output growth

14

would not vary much across time, so little would be gained in terms of the likelihood function value. Consider, in contrast, the non-Markovian regime switching from eq. (7), in which transition probabilities automatically are time-varying. In this model one can have both frequent regime transitions and one-step-ahead forecasts of output growth that vary considerably across time. All it takes in eq. (7) is for the unconditional mean of the latent variable, S^*, to be near zero and for the autoregressive coefficient, θ_1, to be greater than zero. Then, the conditional mean of S^* can differ from zero, causing the one-step-ahead forecasts of output growth to differ from the unconditional mean.

For the model with the informative priors discussed above, the MCMC sampler was run through 40,000 iterations with the first 10,000 iterations discarded to allow the sampler to converge on the posterior distribution. The multi-move Metropolis-Hastings sampler of the latent regime strength variable is efficient enough that this number of iterations ensures that the posterior mean of the latent variable vector is replicated across numerous estimation runs. Quarterly GDP growth data from 1960Q1 to 2003Q4 were used to estimate the non-Markovian regime switching model. Table 2 shows the posterior means and 90 percent probability intervals for the coefficients corresponding to the self-contained model of equation (7), where a lagged dependent variable is the only covariate in the latent state strength equation. In this case, the identification of the model parameters in the latent state equation is tied to the distributional assumption of normality.

Table 2: Coefficient posterior distributions for self-contained model		
	Posterior dist.	Prior
Observation equation		
α_0	0.140	0.10
	(0.044,0.236)	
α_1	0.868	0.80
	(0.778,0.956)	
ϕ	0.186	Flat
	(0.080,0.289)	
Latent regime equation		
λ	0.265	Flat
	(-0.007,0.565)	
θ_1	0.611	0.80
	(0.473,0.745)	
θ_2	0.042	Flat
	(-0.190,0.285)	
Covariance matrix		
σ_1^2	0.617	Flat
	(0.494,0.777)	
ρ	0.160	Flat
	(-0.052,0.348)	
90% prob. interval in parentheses		

The posterior means of the α intercepts in the observation equation move less from the prior values than does the autoregressive coefficient, θ_1, in the latent regime equation. The fact that the model finds that the autoregressive coefficient, θ_1, is centered far from zero in the latent regime equation supports the idea that the index of regime strength, S^*, responds more closely to its own past value than to other variables, such as y_{t-1}. For the self-contained model, the probability interval for the covariance parameter, ρ is not decisively positive, so the evidence in favor of regime endogeneity is not overwhelming. We re-examine the probability interval for this parameter of regime endogeneity below for a specification that is not self-contained and not identified solely through a choice of nonlinear distribution function.

As an alternative specification, we replaced y_{t-1} as a covariate in the latent regime equation (7) with the lagged change in the index of leading indicators. In this case, a unique source of movement in the latent regime strength index, S^*, from the leading indicators helps identify the latent variable, apart from the nonlinear identification from the distribution function. The coefficient on the leading indicators is denoted θ_2, as the leading indicators simply take the place of y_{t-1} in the latent regime equation. Table 3 presents the posterior means of the coefficients for the specification that uses the lagged change in the log of the leading indicators as a predetermined covariate.

Table 3: Coefficient posterior distributions for specification with leading indicator covariate		
	Posterior dist.	Prior
	Observation equation	
α_0	0.130	0.10
	(0.034,0.222)	
α_1	0.885	0.80
	(0.794,0.978)	
ϕ	0.193	Flat
	(0.088,0.298)	
	Latent regime equation	
λ	0.216	Flat
	(-0.027,0.469)	
θ_1	0.547	0.80
	(0.412,0.679)	
θ_2	0.116	Flat
	(0.013,0.225)	
	Covariance matrix	
σ_1^2	0.586	Flat
	(0.478,0.731)	
ρ	0.135	Flat
	(-0.063,0.321)	
90% prob. interval in parentheses		

On the whole, the parameter estimates are quite similar across the two specifications, and this suggests that the identification that hinges on the normality assumption is not

terribly off base. Nevertheless, one difference between the two specifications is that the leading indicators covariate has a 90 percent probability interval that lies only in the positive region. In this case, the leading indicators add some regime forecasting power beyond that brought by the autoregressive term, whereas lagged GDP growth did not. Like the specification without an instrument, the estimated covariance parameter ρ is positive, although its 90 percent probability interval includes zero. This result stands in contrast to the typical Markov switching model which assumes that the unobserved regimes are determined exogenously from the observed data.

Next we compare the posterior means of the latent variable from these two non-Markovian regime switching models with business cycle turning points defined by the National Bureau of Economic Research (NBER).

Matching NBER business cycle turning points

The most interesting output from the non-Markovian regime switching model of GDP growth is the posterior mean of the latent, strength-of-regime indicator, S^*. The results from the self-contained model reported in Table 2 are presented first. Figure 2 shows the posterior mean of the latent regime strength index and how well its crossings of zero match the NBER business cycle turning points. The biggest discrepancies between the sign of the posterior mean of the latent variable and the NBER recession dates occur in the relatively mild 1970 and 2001 recessions. In both cases, the latent regime index dip below zero a bit earlier than the onset of the NBER recession and they also move back above zero a bit before the NBER-dated trough. at the March 1991 NBER trough date. Overall, however, the regime switching model implies switching dates that are very close to NBER turning points throughout the sample. The posterior mean of the latent variable

18

can also serve as a business cycle index, given that it measures the strength of growth rate regimes. For example, 1965, 1972, 1978 and 1984 are periods of pronounced cyclical strength. Similarly, the milder recessions are reflected in the posterior mean of the latent variable as recessions where the the latent variable did not dip as far below zero, such as in 1960, 1970 and 2001. We also calculate posterior means of the regime probabilities, calculated as the percentage of the draws that the latent variable, S_t^*, was above zero. For this measure, a posterior mean probability of 0.5 corresponds closely with NBER turning points, as shown in Figure 3. The priors needed to induce the latent variable to change signs at the business cycle frequency included a prior to keep the two growth rate parameters, α_0 and α_1, sufficiently far apart so as not to reflect fast and moderate growth within economic expansions; the other prior was to ensure a degree of persistence on the latent variable by way of the autoregressive coefficient, θ_1.

The corresponding figures for the specification that used the growth in the index of leading indicators as an instrument in the latent regime strength equation are Figures 4 and 5. The results for the posterior mean of the latent variable and the posterior mean of the regime probabiilities are largely the same as they were in Figures 2 and 3 without an instrument, although the 1980 regime shift starts earlier with the instrument. In addition, the specification that uses the leading indicators also finds a near-recession in 1995, when a recession scare, emanating from a false alarm from the leading indicators among other sources, led the Federal Reserve to cut the federal funds rate three times between July 1995 and January 1996.

Time-varying expected regime durations

With positive serial dependence, the farther the autoregressive latent variable is from

zero (the greater the strength of the current regime), the higher is the expected time before a sign (regime) change. Here we illustrate this feature of the self-contained non-Markovian regime switching model with calculations of time-varying expected durations. Starting with the posterior mean value of the latent variable, S_t^*, and posterior mean values of the parameters, we simulated shock processes for eq. (7) until the sign changed at S_{t+k}^*, where k is the duration of the regime from time t. Note that if we started from the value of $S_t^{*(i)}$ from each iteration i of the MCMC sampler, we would be mixing cases where $S_t^{*(i)}$ was positive and negative. For this reason, we use the posterior mean values as a common starting point. The mean value of k from the simulations was calculated as the expected duration. Figure 6 plots these expected durations. Given the positive intercept, λ, the expected durations are longer on average when S_t^* is above zero (for expansions) than when it is below zero (for recessions). On average, the expected duration in the expansion regime is about three times as long as in the recession regime. This ratio suggests that about 25 percent of the observations will pertain to the recession regime and this figure is not much different from the 21 percent of quarters that the NBER has declared to be recessions. Of course, the expected durations presented here are only in-sample estimates for the purpose of illustrating this feature of the model.

Conclusions

In this article, we present a non-Markovian regime switching model in which the magnitude of the latent variable indexes the time-varying strength of the regime. In our application to regime switching at the business cycle frequency in the growth rate of GDP, we find that the posterior mean of the latent variable looks much like a business cycle index that indicates the degree of cyclical strength or weakness in the economy. Another

useful feature of the non-Markovian model is the time-varying nature of its transition probabilities. For the self-contained model, it is straightforward to calculate the expected duration of the current regime at each observation.

We also demonstrate the straightforward adjustment one can make within the MCMC estimation procedure to allow for the regime process to be correlated with the observable data. Our estimates of GDP growth indicate that the regimes are likely not independent from the observed data. This feature helps regime-switching models confront regime switches where the pressure for a change in regime builds gradually across time.

In terms of methodology, we exploit the fact that it is simple to take expectations of nonlinear regime indicator functions when performing the Kalman filtering as part of a multi-move sampling procedure for the latent regime strength variable. This leads to more natural updates in the Kalman filter equation and better inferences of the latent variable.

REFERENCES

Albert, James H. and Siddhartha Chib (1993), "Bayes Inference via Gibbs Sampling of Autoregressive Time Series, Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics* 11, 1-15.

Carter, C.K. and P. Kohn (1994), "On Gibbs Sampling for State-Space Models," *Biometrica* 81, 541-53.

Chib, Siddhartha and Edward Greenberg (1996), "Markov Chain Monte Carlo Simulation Methods in Econometrics," *Econometric Theory* 12, 409-31.

DeJong, Robert and Tieman Woutersen (2004), "Dynamic Time Series Binary Choice," manuscript Ohio State University.

Eichengreen, B. and M.W. Watson and R.S. Grossman (1985), "Bank Rate Policy under the Interwar Gold Standard," *Economic Journal* 95, 725-45.

Filardo, Andrew J. and Stephen F. Gordon (1998), "Business Cycle Durations," *Journal of Econometrics* 85, 99-123.

Filardo, Andrew J. (1994), "Business-Cycle Phases and Their Transitional Dynamics," *Journal of Business and Economic Statistics* 12, 299-308.

Hamilton, James D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, 357-84.

Horowitz, Joel (1992), "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica* 60, 505-31.

Kim, Chang-Jin, Jeremy Piger and Richard Startz (2003), "Estimation of Markov Regime-Switching Regression Models with Endogenous Switching," Federal Reserve Bank of St. Louis Working Paper 2003-015.

Lam, Pok-Sang (2004), "A Markov Switching Model of GNP Growth with Duration Dependence," *International Economic Review* 45, 175-204.

Sims, Christopher and Tao Zha (2004), "Were There Regime Switches in U.S. Monetary Policy?" manuscript, Princeton University.
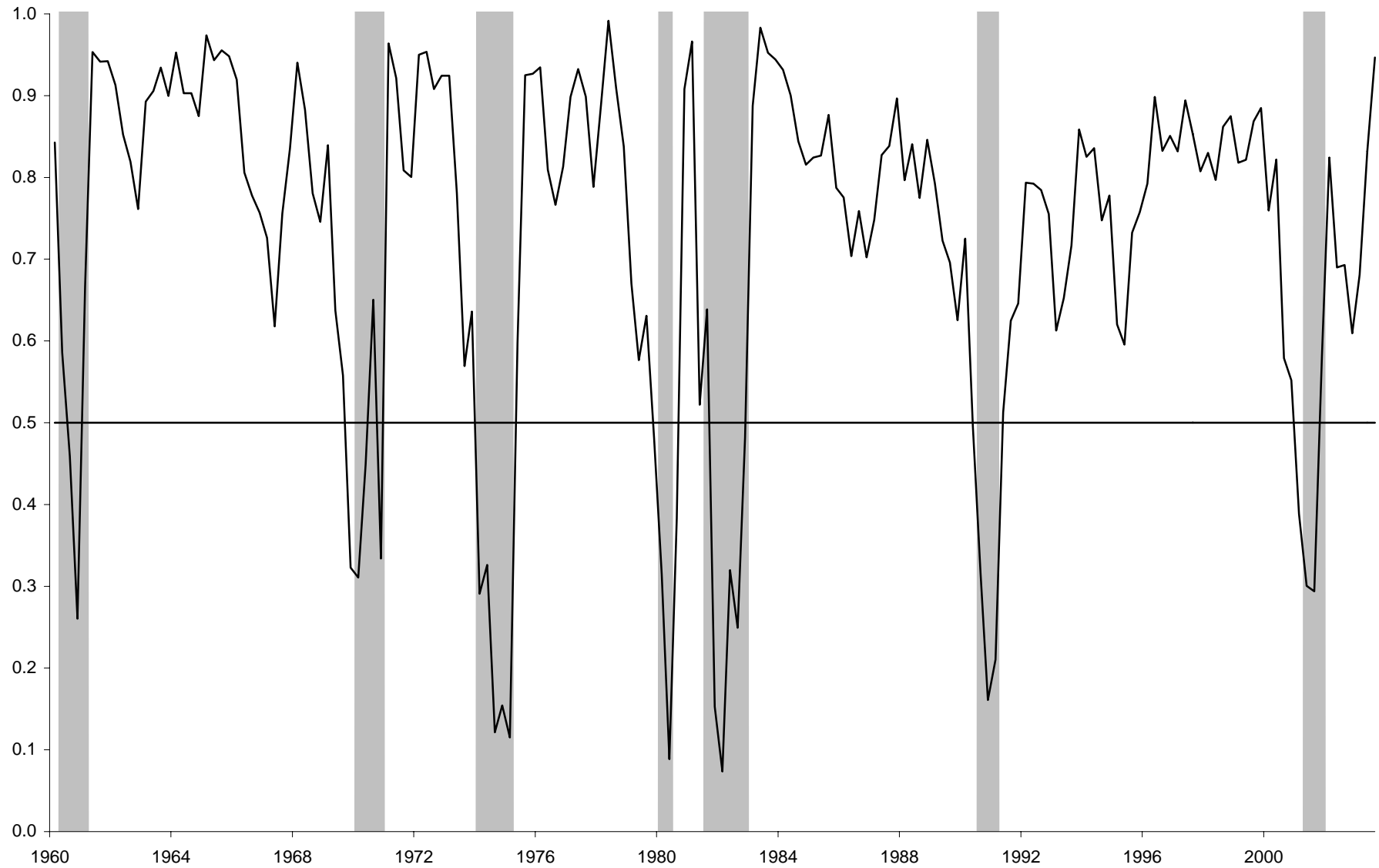
Performance of Sampling the Latent Variable in Monte Carlo
(Deviation from True Value; Every 15th Observation)
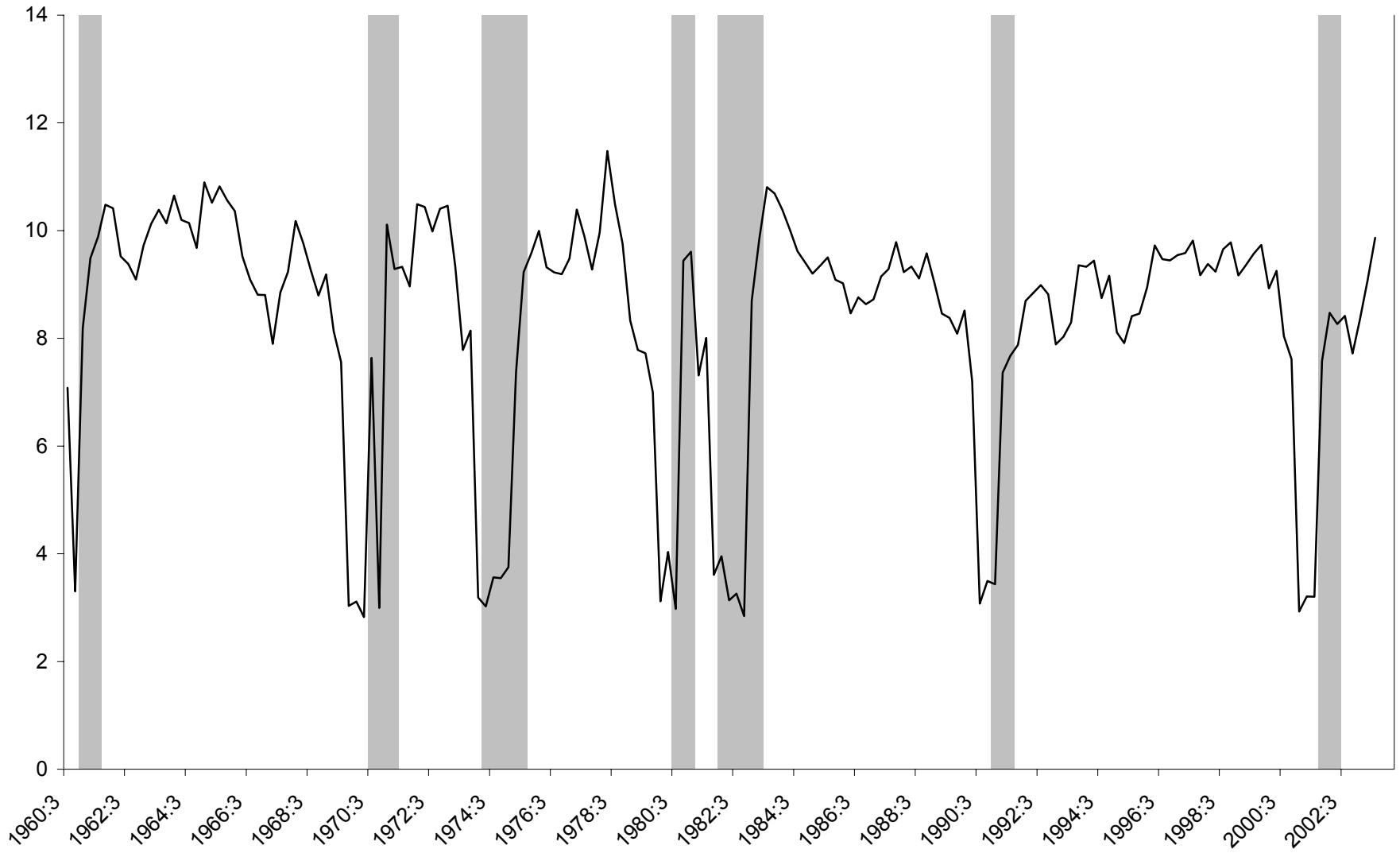

upper 5%   lowest 5%   - mean
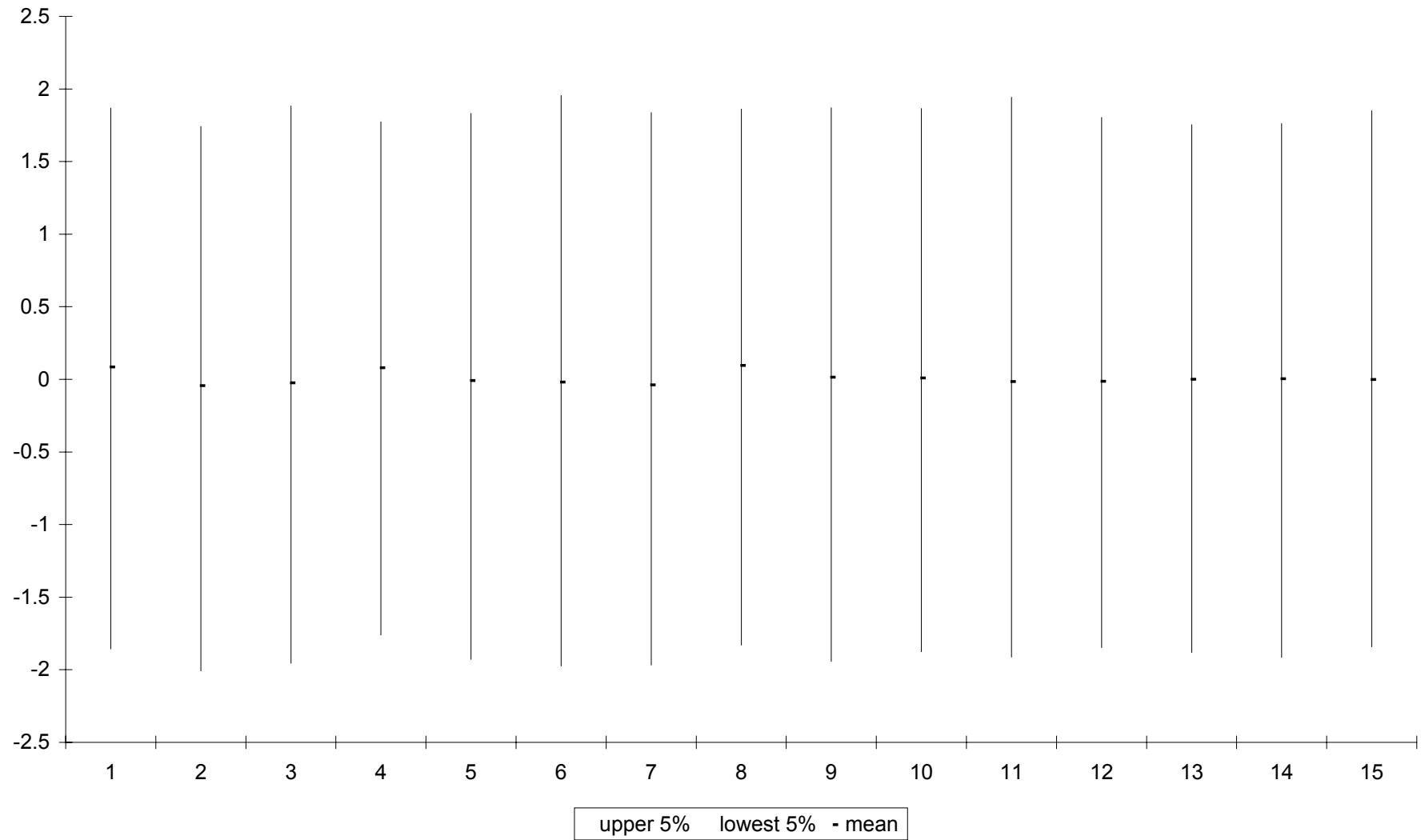
**Figure 2: Posterior Mean of Latent Regime Strength
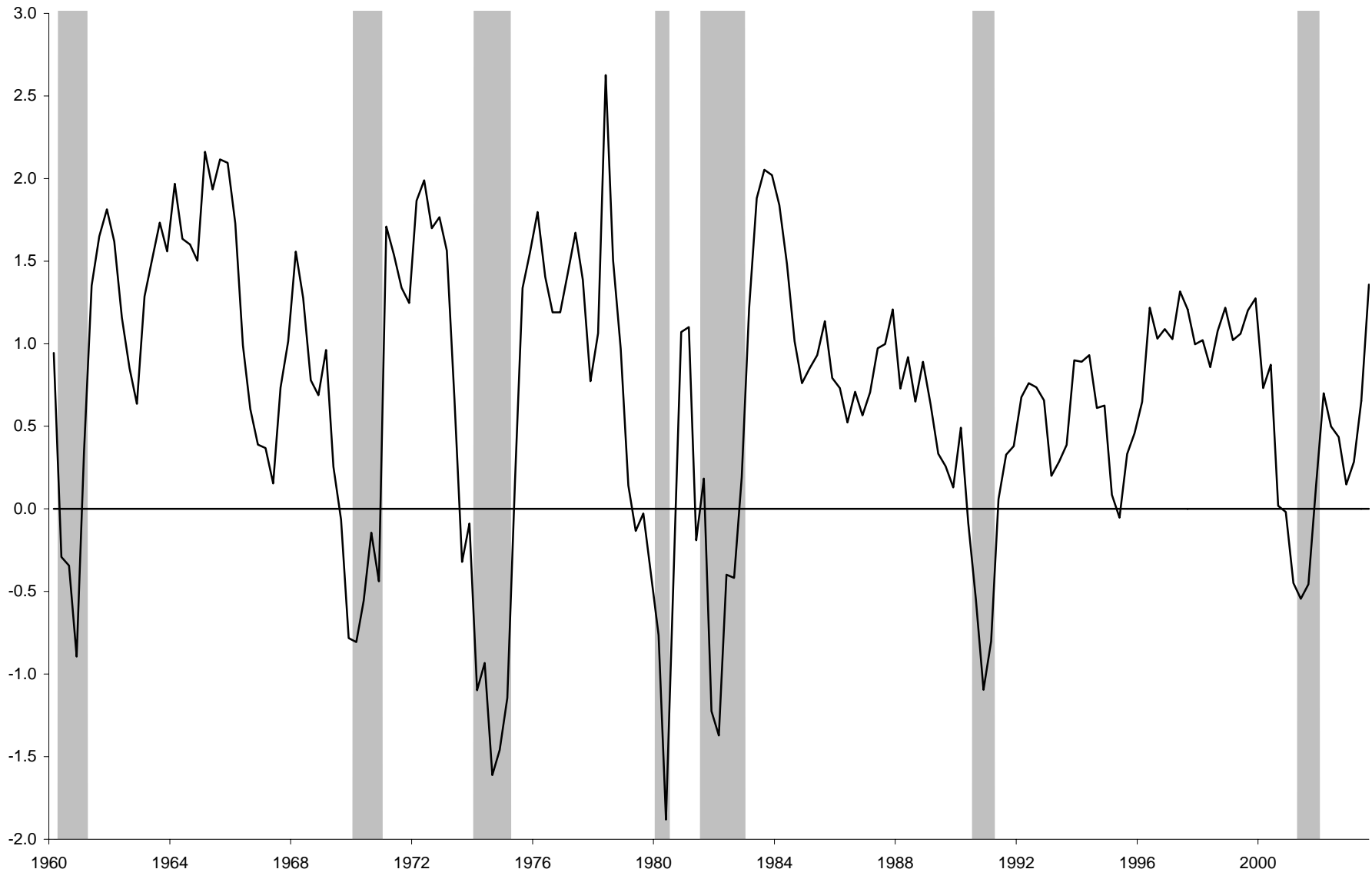from Self-Contained Model (no instrument)**
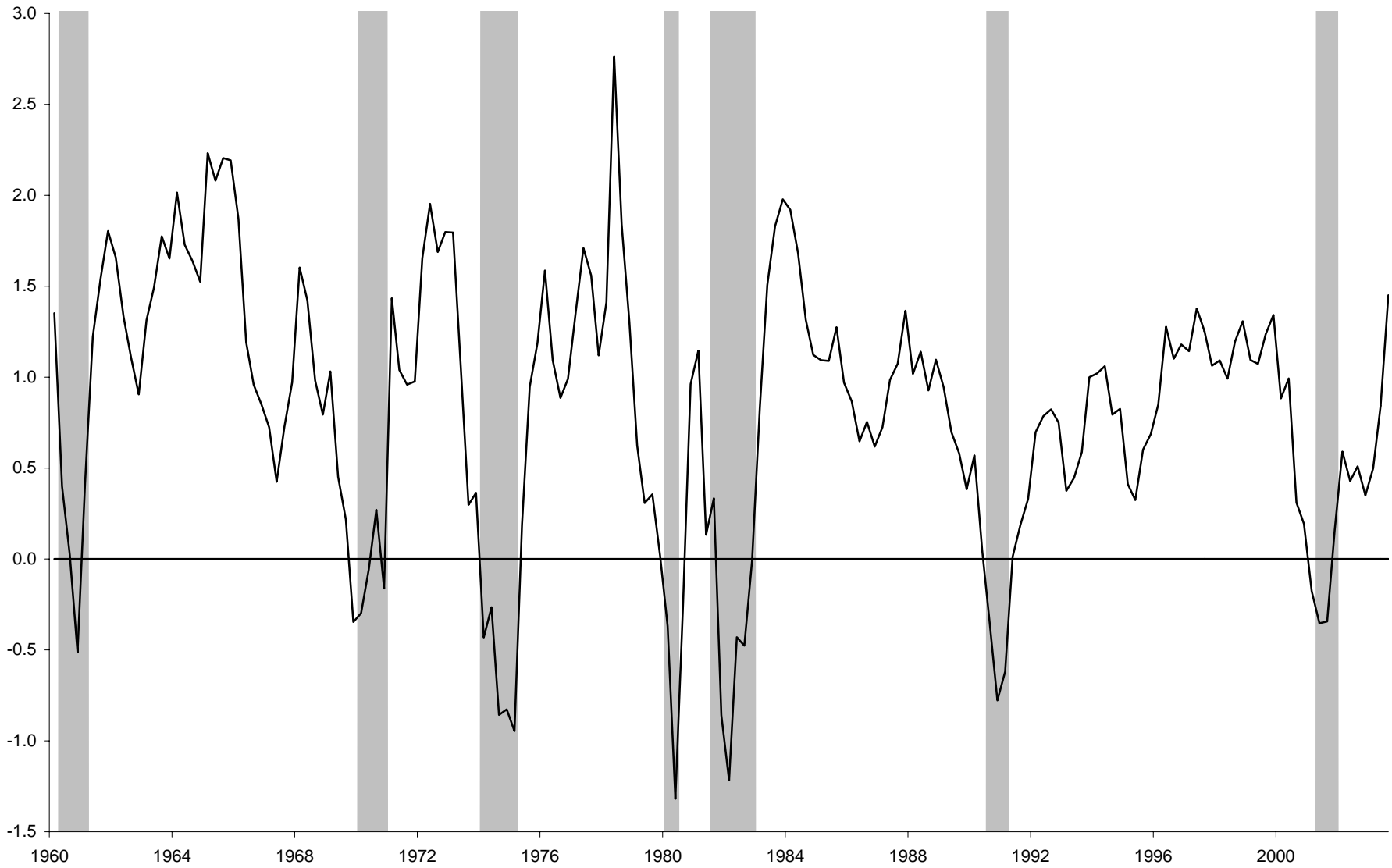
Figure 3: Posterior Mean of the Probablility of the High-Growth Regime from Self-Contained Model (no instrument)

Expected Regime Duration Implied by Self-Contained Model

**Figure 5: Posterior Mean of the Latent Regime Strength
Using Leading Indicators as an Instrument**

Figure 6: Posterior Mean of the Probablility of the High-Growth Regime Using Leading Indicators as an Instrument