Learning and the Great Moderation

Learning and the Great Moderation

James Bullard[*]
Federal Reserve Bank of St. Louis

Aarti Singh[†]
University of Sydney

3 February 2011[‡]

Abstract

We study a stylized theory of the volatility reduction in the U.S. after 1984—the Great Moderation—which attributes part of the stabilization to less volatile shocks and another part to more difficult inference on the part of Bayesian households attempting to learn the latent state of the economy. We use a standard equilibrium business cycle model with technology following an unobserved regime-switching process. After 1984, according to Kim and Nelson (1999a), the variance of U.S. macroeconomic aggregates declined because boom and recession regimes moved closer together, keeping conditional variance unchanged. In our model this makes the signal extraction problem more difficult for Bayesian households, and in response they moderate their behavior, reinforcing the effect of the less volatile stochastic technology and contributing an extra measure of moderation to the economy. We construct example economies in which this learning effect accounts for about 30 percent of a volatility reduction of the magnitude observed in the postwar U.S. data. *Keywords*: Bayesian learning, information, business cycles, regime-switching. *JEL codes*: E3, D8.

1 Introduction

1.1 Overview

The U.S. economy experienced a significant decline in volatility—on the order of fifty percent for many key macroeconomic variables—sometime during the mid-1980s. This phenomenon, sometimes called the *Great Moderation*, has been the subject of a large and expanding literature. The main question in the literature has been the nature and causes of the volatility reduction. In some of the research, better counter-cyclical monetary policy has been promoted as the main contributor to the low volatility outcomes. In other strands, the lower volatility is attributed primarily or entirely to the idea that the shocks buffeting the economy have generally been less frequent and smaller than those from the high volatility 1970s era. In fact, this is probably the leading explanation in the literature to date. Yet, it strains credulity to think that the full amount of the volatility reduction is simply due to smaller shocks. Why should shocks suddenly be 50 percent less volatile?

In this paper, we study a version of the smaller shock story, but one which we think is more credible. In our version, the economy is indeed buffeted by smaller shocks after the mid-1980s, but this lessened volatility is coupled with changed equilibrium behavior of the private sector in response to the smaller shocks. The changed behavior comes from a learning effect which is central to the paper. The learning effect reduces overall volatility of the economy still further in response to the smaller shock volatility. Thus, in our version, the Great Moderation is due partly to less volatile shocks and partly to a learning effect, so that the shocks do not have to account for the entire volatility reduction. Quantifying the magnitude of this effect in an equilibrium setting is the primary purpose of the paper.

1.2 What we do

There have been many attempts to quantify the volatility reduction in the U.S. macroeconomic data. In this paper we follow the regime-switching

approach to this question, as that will facilitate our learning analysis. The regimes can be thought of as expansions and recessions. According to Kim and Nelson (1999a), expansion and recession states moved closer to one another after 1984, but in a way that kept conditional, within-regime variance unchanged. These results imply that recessions and expansions were relatively distinct phases and hence easily distinguishable in the pre-1984 era. In contrast, during the post-1984 era, the two phases were much less distinct.

The Kim and Nelson (1999a) study is a purely empirical exercise. We want to take their core finding as a primitive for our quantitative-theoretic model: Regimes moved closer together, but conditional variance remained constant. The economies we study and compare will all be in the context of this idea.

We assume that the two phases are driven by an unobservable variable, and that economic agents must learn about this variable by observing other macroeconomic data, such as real output. Agents learn about the unobservable state via Bayesian updating. When the two states are closer together, agents find it harder to infer whether the economy is in a recession or in an expansion based on observable data since the two phases of the business cycle are less distinct. Therefore, learning becomes more difficult and leads to an additional change in the behavior of households. In particular, volatility in macroeconomic aggregates will be moderated since the households are more uncertain which regime they are in at any point in time.

We wish to study this phenomenon in a model which can provide a well-known benchmark. Accordingly, we use a simple equilibrium business cycle model in which the level of productivity depends in part on a first-order, two-state Markov process. The complete information version of this model is known to be very close to linear for quantitatively plausible technology shocks, so that a reduction in the variance of the driving shock process translates one-for-one into a reduction in the variance of endogenous variables in equilibrium. The incomplete information, Bayesian learning version of the model is nonlinear. Reductions in driving shock

variance result in more than a one-for-one reduction in the variance of endogenous variables. The difference between what one observes in the complete information case and what one observes in the incomplete information, Bayesian learning case is the learning effect we wish to focus upon.

1.3 Main findings

We begin by establishing that the baseline complete information model with regime-switching behaves nearly identically to standard models in this class under complete information when we use a suitable calibration that keeps driving shock variance and persistence at standard values. We then use the incomplete information, Bayesian learning version of this model as a laboratory to attempt to better understand the learning effect in which we are interested.

We begin by reporting results obtained by allowing unconditional variance to rise as regimes are moved farther apart, keeping conditional variance constant. We compare the resulting volatility of endogenous variables to a complete information model. The complete information model is close to linear, and so the volatility of endogenous variables relative to the volatility of the shock is a constant. For the incomplete information, Bayesian learning economies, endogenous variable volatility rises with the volatility of the shock. This ratio begins to approach the complete information constant for sufficiently high shock variance. Thus the incomplete information economies begin to behave like complete information economies when the two regimes are sufficiently distinct. This is because the inference problem is simplified as the regimes move apart, and thus agent behavior is moderated less.

We then turn to a quantitative assessment of the moderating force in two calibrated incomplete information economies. In these two economies observed volatility in macroeconomic variables is substantially different, with one economy enjoying on the order of 50 percent lower output volatility than the other. This volatility difference is then decomposed into a portion due to lower shock variance and another portion due to more diffi-

cult inference—the learning effect in which we are interested. We find that the learning effect accounts for about 30 percent of the volatility reduction, and the smaller shock portion accounts for about 70 percent. This suggests that learning effects may help account for a substantial fraction of observed volatility reduction in more elaborate economies which can confront more aspects of observed macroeconomic data.

Finally, we turn to consider economies in which the stochastic driving process is estimated via methods similar to those employed by Kim and Nelson (1999a), for 1954:1 to 2004:4 data with 1983:4 as an exogenous break date.[1] We then compute volatility reductions implied by these estimates, and the fraction of the volatility reduction that can be attributed to the learning effect in which we are interested. We find that the total volatility reduction implied by these estimates is about 35 percent for output in our baseline estimated case. This is about two-thirds of the volatility reduction that we observe in the data. Within this reduction, about 43 percent is due to learning, while the other 57 percent is due to the regimes moving closer together. In this empirical section we discuss in more detail the moderation effects as they apply to other variables, mainly consumption, labor hours, and investment. We also include a discussion of serial correlation in these variables associated with the moderation. In general, we think this model is not sufficiently rich to effectively confront the data at this level of detail, but we offer this discussion in an attempt to be as complete as possible and to offer some guidelines for future research on incomplete information economies.

1.4   Recent related literature

Broadly speaking, there are two strands of literature concerning the Great Moderation. One focuses on dating the Great Moderation and the other

---

[1]The calibrated case has the advantage of remaining closer to the standard equilibrium business cycle literature, thus providing a benchmark, while the estimated case has the advantage that the technology process is estimated using a regime-switching process, one of the key features of our stylized model.

looks into the causes that led to it. The dating literature, including Kim and Nelson (1999a), McConnell and Perez-Quiros (2000), and Stock and Watson (2003) typically assumes that the date when the structural break occurred is unknown and then identifies it using either classical or Bayesian methods. According to the other strand, there are three broad causes of the sudden reduction in volatility—better monetary policy, structural change, or luck. Clarida, Gali and Gertler (2000) argued that better monetary policy in the Volker-Greenspan era led to lower volatility, but Sims and Zha (2006) and Primiceri (2005) conclude that switching monetary policy regimes were insufficient to explain the Great Moderation and so favor a version of the luck story. The proponents of the structural change argument mainly emphasize one of two reasons for reduced volatility: a rising share of services in total production, which is typically less volatile than goods sector production (Burns (1960), Moore and Zarnowitz (1986)), and better inventory management (Kahn, McConnell, and Perez-Quiros (2002)).[2] A number of authors, including Ahmed, Levin, and Wilson (2004), Arias, Hansen, and Ohanian (2007), and Stock and Watson (2003) have compared competing hypotheses and concluded that in recent years the U.S. economy has to a large extent simply been hit by smaller shocks.[3]

In the literature, learning has often been used to help explain fluctuations in endogenous macroeconomic variables. In Cagetti, Hansen, Sargent and Williams (2002) agents solve a filtering problem since they are uncertain about the drift of the technology. In Van Nieuwerburgh and Veldkamp (2006) agents solve a problem similar to the one posed in this paper. They use their model to help explain business cycle asymmetries. In their paper learning asymmetries arise due to an endogenously varying rate of information flow. Andolfatto and Gomme (2003) also incorporate learning in a

---

[2]Kim, Nelson, and Piger (2004) argue that the time series evidence does not support the idea that the volatility reduction is driven by sector specific factors.

[3]Owyang, Piger, and Wall (2008) use state-level employment data and document significant heterogeneity in volatility reductions by state. They suggest that the disaggregated data is inconsistent with the inventory management hypothesis or less volatile aggregate shocks, and instead favors the improved monetary policy hypothesis.

dynamic stochastic general equilibrium model. Here agents learn about the monetary policy regime, instead of technology, and learning is used to help explain why real and nominal variables may be highly persistent following a regime change. In an empirical paper, Milani (2007) uses Bayesian methods to estimate the impact of learning in a DSGE New Keynesian model. In his model, recursive learning contributes to the endogenous generation of time-varying volatility similar to that observed in the U.S. postwar period.[4]

Arias, Hansen and Ohanian (2007) employ a standard equilibrium business cycle model as we do, but with complete information. They conclude that the Great Moderation is most likely due to a reduction in the volatility of technology shocks. Our explanation does rely on a reduction in the volatility of technology shocks but that reduction accounts for only a fraction of the moderation according to the model in this paper.

1.5 Organization

In the next section we present our model. In the following section we calibrate and solve the model using perturbation methods. We then report results for a particular calibrated case in order to fix ideas and provide intuition. The subsequent section turns to results based on an estimated regime-switching process for technology. The final section offers some conclusions and suggests directions for future research.

2 Environment

2.1 Overview

We study an incomplete information version of an equilibrium business cycle model. We think of this model as a laboratory to study the effects in which we are interested. Time is discrete and denoted by $t = 0, 1, ...\infty$.

---

[4]Two additional empirical papers, Justiniano and Primiceri (2008) and Fernandez-Villaverde and Rubio-Ramirez (2007), introduce stochastic volatility into DSGE settings, but without learning, and conclude that volatilities have changed substantially during the sample period.

The economy consists of an infinitely-lived representative household that derives utility from consumption of goods and leisure. Aggregate output is produced by competitive firms that use labor and capital.

2.2 Households

The representative household is endowed with 1 unit of time each period which it must divide between labor, ℓ_t, and leisure, $(1 - \ell_t)$. In addition, the household owns an initial stock of capital k_0 which it rents to firms and may augment through investment, i_t. Household utility is defined over a stochastic sequence of consumption c_t and leisure $(1 - \ell_t)$ such that

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, \ell_t), \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, E_0 is the conditional expectations operator and period utility function u is given by

$$u(c_t, \ell_t) = \frac{[c_t^{\theta}(1 - \ell_t)^{1-\theta}]^{1-\tau}}{1 - \tau}. \tag{2}$$

The parameter τ governs the elasticity of intertemporal substitution for bundles of consumption and leisure, and θ controls the intratemporal elasticity of substitution between consumption and leisure. At the end of each period t the household receives wage income and interest income. Thus the household's end-of-period budget constraint is[5]

$$c_t + i_t = w_t \ell_t + r_t k_t, \tag{3}$$

where i_t is investment, w_t is the wage rate, and r_t is the interest rate. The law of motion for the capital stock is given by

$$k_{t+1} = (1 - \delta)k_t + i_t, \tag{4}$$

where δ is the net depreciation rate of capital.

---

[5]We stress "end of period" since in the begining of the period the agent has only expectations about the wage and the interest rate.

2.3 Firms

Competitive firms produce output y_t according to the constant returns to scale technology

$$y_t = e^{z_t} f(k_t, \ell_t) = e^{z_t} k_t^\alpha \ell_t^{1-\alpha}, \tag{5}$$

where k_t is aggregate capital stock, ℓ_t is the aggregate labor input and z_t is a stochastic process representing the level of technology relative to a balanced growth trend.

2.4 Shock process

We assume that the level of technology is dependent on a latent variable. Accordingly, we let z_t follow the stochastic process[6]

$$z_t = (a_H + a_L)(s_t + \varsigma \eta_t) - a_L, \tag{6}$$

with

$$z_t = \begin{cases} a_H + (a_H + a_L)\varsigma \eta_t & \text{if } s_t = 1 \\ -a_L + (a_H + a_L)\varsigma \eta_t & \text{if } s_t = 0 \end{cases}, \tag{7}$$

where $a_H \geq 0$, $a_L \geq 0$, $\eta_t \sim i.i.d.$ $N(0,1)$, and $\varsigma > 0$ is a weighting parameter. The variable s_t is the latent state of the economy where $s_t = 0$ denotes a "recession" state, and $s_t = 1$ denotes an "expansion" state. We assume that s_t follows a first-order Markov process with transition probabilities given by

$$\Pi = \begin{pmatrix} q & 1-q \\ 1-p & p \end{pmatrix}, \tag{8}$$

where $q = \Pr(s_t = 0 | s_{t-1} = 0)$ and $p = \Pr(s_t = 1 | s_{t-1} = 1)$. Hamilton (1989) shows that the stochastic process for s_t is stationary and has an $AR(1)$ specification such that

$$s_t = \lambda_0 + \lambda_1 s_{t-1} + v_t, \tag{9}$$

where $\lambda_0 = (1-q)$, $\lambda_1 = (p + q - 1)$, and v_t has the following conditional probability distribution: If $s_{t-1} = 1$, $v_t = (1-p)$ with probability p and

---

[6] As we discuss below, the process is written in this form to facilitate our use of perturbation methods as implemented by Aruoba, et al., (2006).

$v_t = -p$ with probability $(1 - p)$; and, $v_t = -(1 - q)$ with probability q and $v_t = q$ with probability $(1 - q)$ conditional on $s_{t-1} = 0$. Thus

$$z_t = \xi_0 + \xi_1 z_{t-1} + \sigma \epsilon_t, \tag{10}$$

where $\xi_0 = (a_H + a_L) \lambda_0 + \lambda_1 a_L - a_L$, $\xi_1 = \lambda_1$, $\sigma = (a_H + a_L)$, and

$$\epsilon_t = v_t + \varsigma \eta_t - \lambda_1 \varsigma \eta_{t-1}. \tag{11}$$

The stochastic process for z_t, equation (10), has the same $AR(1)$ form as in a standard equilibrium business cycle model even though we have incorporated regime-switching. In our quantitative work, we use σ as the perturbation parameter in order to approximate a second-order solution to the equilibrium of the economy. We sometimes call σ the "regime distance" as it measures the distance between the conditional expectation of the level of technology relative to trend, z, in the high state, a_H, and the low state, $-a_L$. The distribution of ϵ is nonstandard, being the sum of discrete and continuous random variables. Since η is *i.i.d.*, v and η are uncorrelated. The mean of ϵ_t is zero and the variance is given by

$$\sigma_\epsilon^2 = p(1 - p) \frac{\lambda_0}{1 - \lambda_1} + q(1 - q) \left(1 - \frac{\lambda_0}{1 - \lambda_1}\right) + \varsigma^2 \left(1 + \lambda_1^2\right). \tag{12}$$

We draw from this distribution when simulating the model. The variance is in part a function of ς, which will play a role in the analysis below.

2.5 Information structure

2.5.1 Overview

In any period t, the agent enters the period with an expectation of the level of technology, z_t^e. The latent state, s_t, as well as the two shocks η_t and v_t, are all unobservable by the agent. First, households and firms make decisions. Next, shocks are realized and output is produced. We let actual consumption equal planned consumption and require investment to absorb any difference between expected output and actual output. At the end of the period, the level of technology z_t can be inferred based on the amount

of inputs used and the realized output, since $z_t = \log y_t - \log k_t^\alpha \ell_t^{1-\alpha}$. The agents use z_t in part to calculate next period's expected latent state, s_{t+1}^e, using Bayes' rule, and then the expected level of technology for the next period, z_{t+1}^e. Period t ends and the agent enters the next period with z_{t+1}^e. The details of these calculations are given below. Given this timing, the information available to the agent at the time decisions are made is $F_t = \{y^{t-1}, c^{t-1}, z^{t-1}, i^{t-1}, k^t, \ell^{t-1}, w^{t-1}, r^{t-1}\}$. Here $h^t = \{h_0, h_1, ..., h_t\}$ represents the history of any series h.

2.5.2 Expectations

As shown in the appendix, the current expected state is given by

$$s_t^e = b_t(1 - q) + (1 - b_t)p, \tag{13}$$

where $b_t = P(s_{t-1} = 0 | F_t)$ and the expected level of technology at date t is given by

$$z_t^e = (a_H + a_L)s_t^e + (-a_L). \tag{14}$$

Equivalently

$$z_t^e = [b_t(1 - q) + (1 - b_t)p]\, a_H - [b_t q + (1 - b_t)(1 - p)]\, a_L. \tag{15}$$

We stress that the expectation of the level of technology can be written in a recursive way. First, solve equation (15) for b_t to obtain

$$b_t = \frac{(a_H + a_L)\, p - a_L - z_t^e}{(a_H + a_L)\,(p + q - 1)}. \tag{16}$$

Also from equation (15), next period's value of z^e is

$$z_{t+1}^e = [b_{t+1}(1 - q) + (1 - b_{t+1})p]\, a_H - [b_{t+1}q + (1 - b_{t+1})(1 - p)]\, a_L. \tag{17}$$

The value of b_{t+1} in this equation can be written in terms of updated $(t + 1)$ values of g_L and g_H defined in equations (36) and (37) in the appendix, which will depend on b_t, and, through the definitions of the conditional

densities (38), (39), (40), and (41) in the appendix, on z_t as well. Using equation (16) to eliminate b_t we conclude that we can write

$$z_{t+1}^e = f\left(z_t^e, z_t\right) \tag{18}$$

where f is a complicated function of z_t^e and z_t. The fact that z_{t+1}^e has a recursive aspect plays a substantive role in some of our findings below. When the agent infers a value for z_t at the end of the period, that value is not the only input into next period's expected value, as z_t^e also plays a role.

2.6 The household's problem

The household's decision problem is to choose a sequence of $\{c_t, \ell_t\}$ for $t \geq 0$ that maximizes (1) subject to (3) and (4) given a stochastic process for $\{w_t, r_t\}$ for $t \geq 0$, interiority constraints $c_t \geq 0$, $0 \leq \ell_t \leq 1$, and given k_0. Expectations are formed rationally given the assumed information structure.

Assuming an interior solution, the optimality conditions imply that

$$\frac{u_\ell(c_t, \ell_t)}{u_c(c_t, \ell_t)} = w_t \tag{19}$$

Before the shocks are realized, households make their consumption and labor decisions. In this model investment is a residual and absorbs unexpected shocks to income. The Euler equation is

$$u_c(c_t, \ell_t) = \beta E_t \left[u_c(c_{t+1}, \ell_{t+1})(r_{t+1} + (1 - \delta)) \right]. \tag{20}$$

2.7 The firm's problem

Firms produce a final good by choosing capital k_t and labor ℓ_t such that they maximize their expected profits. The firms period t problem is then

$$\max_{k_t, \ell_t} E_t[e^{z_t} k_t^\alpha \ell_t^{1-\alpha} - w_t \ell_t - r_t k_t] \ \forall t. \tag{21}$$

The first order conditions for the firm are

$$r_t = E_t[e^{z_t} f_k(k_t, \ell_t)] \tag{22}$$

11

$$w_t = E_t[e^{z_t} f_\ell(k_t, \ell_t)] \tag{23}$$

These conditions differ from the standard condition because technology level z_t is not observable when decisions are made.

2.8   Second-order approximation

We follow Van Nieuwerburgh and Veldkamp (2006) and study a passive learning problem.[7] As a result, the timing and informational constraints faced by the planner are the same as in the decentralized economy, and the competitive equilibrium and the planning problem are equivalent. The planner's problem is to maximize household utility (1) subject to the resource constraint

$$c_t + k_{t+1} = e^{z_t} k_t^\alpha \ell_t^{1-\alpha} + (1 - \delta) k_t \tag{24}$$

and the evolution of beliefs given by equations (15) and (42) in the appendix. The solution to this problem is characterized by (19), (20), (24), and the exogenous stochastic process (10).

The perturbation methods we use are standard and are described in Aruoba, Fernandez-Villaverde, and Rubio-Ramirez (2006). To solve the problem we find three decision rules, one each for consumption, labor supply, and next period's capital, as a function of the two states (k, z^e) and a perturbation parameter σ. Our regime distance parameter $\sigma = a_H + a_L$ plays the role of σ in Aruoba, et al. (2006).

The core of the perturbation method is to approximate the decision rules by a Taylor series expansion at the deterministic steady state of the model, which is characterized by $\sigma = 0$.[8] For instance, the second-order

---

[7]The planner does not take into account the effect of consumption and labor choices on the evolution of beliefs.

[8]Auroba, et al., (2006) report that higher-order perturbation methods perform better than simple linear approximations when the model is characterized by significant non-linearities. They focus primarily on non-linearities arising due to high shock variance and high risk aversion. In our analysis the non-linearities come about because of the Bayesian learning mechanism. By using a second order approximation our objective is to avoid potential numerical approximation errors due to these non-linearities.

Taylor series approximation to the consumption policy function is

$$
\begin{aligned}
c\left(k, z^e, \sigma\right) \;=\; & c_{ss} + c_k \left(k - k_{ss}\right) + c_{z^e} \left(z^e - z^e_{ss}\right) + c_\sigma \left(\sigma - \sigma_{ss}\right) \\
& + \frac{1}{2} c_{k,k} \left(k - k_{ss}\right)^2 + \frac{1}{2} c_{k,z^e} \left(k - k_{ss}\right) \left(z^e - z^e_{ss}\right) + \ldots \\
& + \frac{1}{2} c_{\sigma,z^e} \left(\sigma - \sigma_{ss}\right) \left(z^e - z^e_{ss}\right) + \frac{1}{2} c_{\sigma,\sigma} \left(\sigma - \sigma_{ss}\right)^2 ,
\end{aligned}
$$

where x_{ss} is the steady state value of a variable (actually zero for z^e_{ss} and σ_{ss}), c_i is the first partial derivative with respect to i, and $c_{i,j}$ is the cross partial with respect to i and then j, and all derivatives are evaluated at the steady state. The program we use calculates analytical derivatives and evaluates them to obtain numerical coefficients c_i and $c_{i,j}$ as well as analogous coefficients for the policy functions governing labor and next period's capital.

3 Learning effects

3.1 Calibration

In this section we follow the equilibrium business cycle literature and calibrate the model at a quarterly frequency. This helps to give us a clear benchmark with which to compare and understand our results.

For the calibration, we remain as standard as possible. The discount factor is set to $\beta = 0.9896$, the elasticity of intertemporal substitution is set to $\tau = 2$, and $\theta = 0.357$ is chosen such that labor supply is 31 percent of discretionary time in the steady state. We set $\alpha = 0.4$ to match the capital share of national income and the net depreciation rate is set to $\delta = 0.0196$. This is also the benchmark calibration used by Aruoba, et al., (2006). Apart from the parameters commonly used in the business cycle literature, there are some additional parameters that capture the regime-switching process. In particular, the parameters a_L and a_H reflect the level of technology z_t relative to trend in recessions and expansions respectively. We also have the transition probabilities p and q as well as a weighting parameter ς.

To obtain a baseline complete information economy that can be compared to the benchmark calibration of Aruoba, et al., (2006), (AFR), we use equation (10), reproduced here for convenience

$$z_t = \xi_0 + \xi_1 z_{t-1} + \sigma \epsilon_t. \tag{25}$$

We wish to choose the values of p, q, a_H, a_L and ς such that $\xi_0 = 0$, $\xi_1 = 0.95$, $\sigma = 0.007$, and $\sigma_\epsilon^2 = 1$, the standard equilibrium business cycle values and the ones used in the benchmark calibration of AFR. To remain comparable to AFR, we would like ϵ_t to be close to a standard normal random variable, with $\sigma = a_H + a_L = 0.007$. To meet this latter requirement, we choose symmetric regimes by setting $a_H = a_L = 0.0035$. Since

$$\xi_1 = \lambda_1 = (p + q - 1),$$

we set $p = q = 0.975$, yielding $\xi_1 = 0.95$. These values imply $\xi_0 = (a_H + a_L)\lambda_0 + \lambda_1 a_L - a_L = 0$. This leaves the mean and variance of ϵ_t. The mean is zero, but to get the variance

$$\sigma_\epsilon^2 = p(1-p)\frac{\lambda_0}{1-\lambda_1} + q(1-q)\left(1 - \frac{\lambda_0}{1-\lambda_1}\right) + \varsigma^2\left(1 + \lambda_1^2\right)$$

equal to one, we set the remaining parameter $\varsigma = 0.719$.[9] Thus the unconditional standard deviation of the shock process, $\sigma\sigma_\epsilon$, is 0.007 as desired, and the conditional standard deviation, $\sigma\varsigma\sigma_\eta = \sigma\varsigma$ is 0.005 (since $\sigma_\eta = 1$ by assumption). For this calibration, the nonstochastic steady state values are given by: $z_{ss}^e = z_{ss} = 0$, $k_{ss} = 23.14$, $c_{ss} = 1.288$, $\ell_{ss} = 0.311$, and $y_{ss} = 1.742$.

3.2 A complete information comparison

In the calibration of the baseline complete information economy, we choose parameters for the regime-switching process such that the economy is as close as possible to a standard equilibrium business cycle model. We now investigate whether the equilibrium of this economy is comparable to the

---

[9]We chose the positive value for ς that met this requirement.

TABLE 1.
Benchmark comparison.

Variable	AFR	Model
Output	1.2418	1.2398
Consumption	0.4335	0.4323
Hours	0.5714	0.5706
Investment	3.6005	3.5953

Table 1: A comparison of standard deviations of key endogenous variables for a standard equilibrium business cycle model (Auroba, et al., (2006), or AFR) and the complete information version of the present model with regime-switching, calibrated to mimic the standard case.

equilibrium of a standard model. Table 1 shows that the baseline complete information model with regime-switching delivers results almost identical to a standard equilibrium business cycle model, that is, the same results as Aruoba, et al, (2006).[10] Since we use perturbation methods to solve our model, it seems natural then to compare our model with Aruoba, et al., (2006).

This shows that despite the addition of regime-switching, the complete information economy calibrated to look like the standard case delivers results very similar to the standard case. We now turn to incomplete information economies with Bayesian learning.

3.3 Incomplete approaches complete information

Our intuition is that, keeping conditional (within regime) variance constant, regimes which are closer together pose a more difficult inference problem for agents. The agents then take actions which are not as extreme as they would be under complete information. The result is a moderating force in the economy.

Our model is designed to allow us to move regimes closer together

---

[10]In Table 1, for both cases, each simulation has 200 observations. For each simulation we compute the standard deviations for percentage deviations from Hodrick-Prescott filter with $\lambda = 1600$ and then average over 250 simulations.

while keeping conditional variance constant. In particular, we can change regime distance $\sigma = a_H + a_L$, and this will change the technology shock volatility $\sigma \epsilon$. At the same time, we can hold conditional (within regime) variance constant using the parameter ς: The conditional standard deviation is just $\sigma \varsigma$. This means that the economies we compare in this subsection will have substantially different unconditional variances, different levels of volatility coming directly from the driving shock process in the economy. We then need a benchmark against which we can compare these economies. The benchmark we choose is the counterpart, complete information version of these economies.

Standard equilibrium business cycle models are known to be very close to linear for quantitatively plausible technology shocks when there is complete information. In these economies, the standard deviation of key endogenous variables increases one-for-one with increases in $\sigma \sigma_\epsilon$. We expect that our regime-switching model with complete information is also very close to linear with respect to the unconditional standard deviation, $\sigma \sigma_\epsilon$. The only difference between this case and the economies we study is the addition of incomplete information and Bayesian learning. The latter economies are nonlinear, so that the standard deviation of key endogenous variables no longer increases one-for-one with increases in $\sigma \sigma_\epsilon$. By comparing the complete and incomplete information versions of the same economies, we can infer the size of the learning effect in which we are interested. In addition, we expect the inference problem to become less severe as regimes move farther apart. The economies with learning should begin to look more like complete information economies as regimes become more distinct.

In Figure 1 we plot the unconditional standard deviation on the horizontal axis. On the vertical axis, we plot the standard deviation of output relative to the unconditional standard deviation. In our version of the standard "RBC" equilibrium business cycle model (no regime-switching, complete information), the standard deviation of output is 1.24 percent and the standard deviation of the shock is 0.7 percent. Thus the ratio of standard

deviation of output relative to the standard deviation of the shock process is 1.77. Because of linearity, this ratio does not change as the unconditional variance of the shock increases. This is depicted by the horizontal dotted line in Figure 1. We also know that our complete information model with regime-switching delivers results close to the standard equilibrium business cycle model for certain parameter values (see Table 1). Thus we expect the ratio of standard deviations of output and shocks to be constant for this case as well. This turns out to be verified in the Figure, as the solid line indicates only minor deviations from the standard equilibrium business cycle model.[11]

The linear relationship between $\sigma\sigma_\epsilon$ and the standard deviation of endogenous variables breaks down in a model with incomplete information. When the states become less distinct, moving from the right to the left along the horizontal axis in Figure 1, the agents have to learn about the state of the economy and the learning effect moderates the behavior of all endogenous variables. But when states are distinct, toward the right in the Figure, the standard deviation of output rises more than one-for-one. Agents are more able to discern the true state when the states are more distinct.

Figure 1 shows that learning has a pronounced effect on private sector equilibrium behavior. Moreover, it shows that the learning effect becomes larger as regimes move closer together, keeping conditional variance unchanged. This makes sense as the inference problem becomes more difficult for agents. The agents base behavior in part on the expected regime, which, because of increased confusion, more often takes on intermediate values instead of extreme values. This leads the agents to take actions midway between the ones they would take if they were sure they were in one regime or the other. This provides a clear moderating force in the economy above and beyond the reduction in unconditional variance. We now turn to a quantitative assessment of the size of this moderating force.

---

[11]Each point in this figure is computed by simulating 200 quarters for the given economy, and averaging results over 250 such economies. We calculate 13 such points and connect them for each line in the figure.
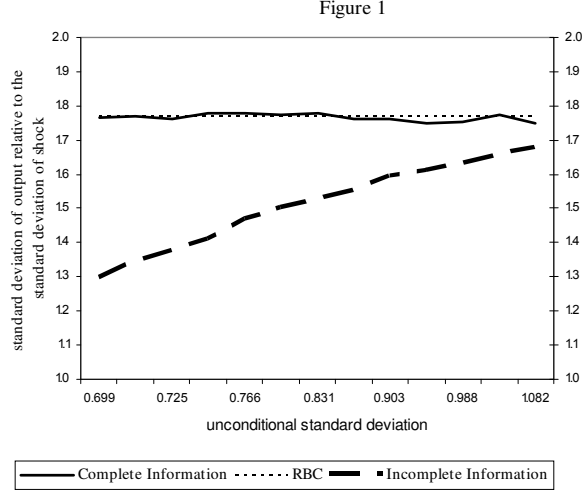
Figure 1: The complete information economy, like the RBC model, has volatility which is proportional to the volatility of the shock. This is indicated by the horizontal line in the Figure. In the incomplete information economy, this is no longer true because of the inference problem. This problem becomes less severe moving to the right in the Figure, and the incomplete information case approaches the complete information case.

3.4 Comparing economies with high and low volatility

The empirical literature on the Great Moderation, including Kim and Nelson (1999a), McConnell and Perez-Quiros (2000), and Stock and Watson (2003), has documented the large decline in output volatility after 1984. As an example, we calculated the Hodrick-Prescott-filtered standard deviation of U.S. output for 1954-1983 and 1984-2004. These values are 1.92 and 0.95, and so the volatility reduction by this measure is $0.95/1.92 \approx 0.50$. In this subsection we want to choose parameters so as to compare two economies across which the cyclical component of output endogenously exhibits a volatility reduction of this magnitude. From there, we want to decompose the sources of the reduction into a portion due to reduced volatility of the shock and another portion due to the learning effect. The main idea is to un-

18

TABLE 2. MODERATION IN CALIBRATED ECONOMIES.

	High volatility economy	Low volatility economy
	Parameter values	
$a_H = a_L$	0.0265	0.0025
$p = q$	0.975	0.975
$\sigma\sigma_\epsilon$	0.0108	0.007
$\sigma\varsigma$	0.005	0.005
	Volatility, in percent standard deviation	
Output	1.816	0.908
Consumption	0.390	0.138
Hours	1.141	0.082
Investment	6.504	3.487
$\frac{1}{T}\sum(s^e - s)^2$	0.125	0.415

Table 2: Comparison of the business cycle volatility in the high and low volatility incomplete information economies. The volatility reduction in the cyclical component of output is about 50 percent, but the volatility reduction in the unconditional variance is only 35 percent. Learning accounts for on the order of 30 percent of the volatility reduction in output.

derstand whether this learning effect could be a quantitatively significant part of an output moderation of this magnitude in a general equilibrium setting.

For this purpose, we set $a_H = a_L = 0.0265$ in the high volatility economy and $a_H = a_L = 0.0025$ in the low volatility economy. This implies $\sigma = 0.053$ in the former case and $\sigma = 0.005$ in the latter case. We again choose ς to keep the conditional standard deviation $\sigma\varsigma$ constant at 0.005. These parameter choices imply that $\sigma\sigma_\epsilon$, the unconditional standard deviation of the productivity shock, is 1.08 percent in the high volatility economy and 0.7 percent in the low volatility economy. We view these as plausible values. We set $p = q = 0.975$, so that $\xi_1 = p + q - 1 = 0.95$ as is standard in the equilibrium business cycle literature. These parameter choices are described in the top panel of Table 2. With these parameter values, the endogenous output standard deviation in the high volatility economy is 1.816, whereas the corresponding standard deviation in the low volatil-

ity economy is 0.908, a reduction of 50 percent. Moreover, all endogenous variables are considerably less volatile.[12] This is documented in the lower panel of Table 2. Our measure of confusion is given in the last line of Table 2. This measure increases substantially as the economy becomes less volatile suggesting that the inference problem becomes more severe in the low volatility economy.[13]

If these were complete information economies, the volatility reduction would be proportional to the decline in the unconditional standard deviation of the productivity shock ϵ_t. If that was the case, endogenous variables in the low volatility economy would be about 65 percent as volatile $\left(\frac{0.007}{0.0108}\right)$ as those in the high volatility economy—this would be a volatility reduction of 35 percent. The actual output volatility reduction is 50 percent, and the extra 15 percentage points of output volatility reduction can be attributed to the learning effect described in the previous subsection. Thus we conclude that for these two economies, the luck part of the output volatility reduction accounts for 35/50 or 70 percent of the total, and the learning effect accounts for 15/50 or 30 percent.

We think this calculation, while far from definitive, clearly demonstrates that learning could play a substantial role in the observed volatility reduction in the U.S. economy, with a contribution that may have been on the order of 30 percent of the total. This is fairly substantial, and it suggests that it may be fruitful to analyze the hypothesis of this paper in more elab-

---

[12]Instead of investment, if consumption is residual, as in Van Nieuwerburgh and Veldkamp (2006), we find that learning still plays a key role in explaining the reduction in output volatility observed in recent years. The main implication of this assumption, however, is on the business cycle volatility of consumption and investment: cyclical component of consumption becomes more volatile while the volatility of investment falls. For example in the baseline calibrated case, the standard deviations of the cyclical component of consumption and investment in the pre-moderation period are 2.0 and 1.8 respectively. In the post-moderation period these volatilities fall to 1.2 and 0.14. The level of consumption and investment volatilities under this assumption are therefore inconsistent with what we observe in the data.

[13]See Campbell (2007) for a discussion of the increased magnitude of forecast errors in the post-moderation era among professional forecasters. One might also view the well-documented increase in lags in business cycle dating in the post-moderation era as an indication of increased confusion between boom and recession states.

orate models which can confront the data on more dimensions.

3.5 Understanding the learning mechanism

3.5.1 Complete information case

One way to understand the effect of incomplete information and Bayesian learning is to first consider the optimal decision rules for the baseline complete information case and then compare them with the corresponding incomplete information economies. For the complete information economies, the state variables are k and z. To simplify our discussion here, we will focus on the first-order terms.[14] In the complete information case, the decision rules for consumption, labor hours, and next period's capital are given as

$$c - c_{ss} = 0.03\,(k - k_{ss}) + 0.60\,(z - z_{ss}) + ... \tag{26}$$

$$l - l_{ss} = -0.002\,(k - k_{ss}) + 0.20\,(z - z_{ss}) + ... \tag{27}$$

$$k' - k_{ss} = 0.97\,(k - k_{ss}) + 1.80\,(z - z_{ss}) + ... \tag{28}$$

Not surprisingly, these rules are almost identical for the high and low volatility economies.[15] Given this, why does the volatility of the endogenous variables change as we move from high to low volatility economies? This is primarily due to changes in the volatility of $(k - k_{ss})$ and $(z - z_{ss})$. In the complete information case, the standard deviation of $(k - k_{ss})$ and $(z - z_{ss})$ decreases in proportion to the decrease in the standard deviation of the underlying shock. As a result, the relative standard deviations of $(k - k_{ss})$ and $(z - z_{ss})$ are almost identical across the two economies, see the top panel of Table 3. Therefore, in the complete information case,

---

[14]Some of the higher order terms are non-zero but they are relatively small and hence not very critical for understanding the learning mechanism. For example, the policy function for consumption in the high volatility economy is $c - c_{ss} = 0.03\,(k - k_{ss}) + 0.60\,(z^e - z_{ss}^e) - 0.0005\,(k - k_{ss})^2 + 0.47\,(z^e - z_{ss}^e)^2 + 0.008\,(k - k_{ss})\,(z^e - z_{ss}^e) - 0.027\,(\sigma - \sigma_{ss})^2$.

[15]Note that the only difference between the two economies is the driving shock process. The coefficient that differs across the two economies in both cases—complete information and incomplete information—is the second order coefficient on the perturbation parameter. However, this coefficient is relatively small to impact the results in any significant way.

TABLE 3. KEY ELEMENTS OF THE DECISION RULE.

Relative standard deviation

	High volatility economy	Low volatility economy
Complete information case		
$k - k_{ss}$	88.26	85.78
$z - z_{ss}$	2.82	2.83
Incomplete information case		
$k - k_{ss}$	104.45	85.58
$z^e - z^e_{ss}$	2.15	0.30

Table 3: Relative volatility of the key elements of the decision rule in the baseline calibrated case. To compute the relative standard deviation, in each case we divide the actual standard deviation by the standard deviation of the technology process, 1.08 percent in the high volatility economy and 0.7 percent in the low volatility economy of our baseline calibrated case.

the volatility of key endogenous variables decreases one-for-one with decreases in $\sigma\sigma_\epsilon$, as noted earlier in section 3.3.

3.5.2   Incomplete information case

The following equations describe the decision rules for the incomplete information economy where the state variables are k and z^e :

$$c - c_{ss} = 0.03\,(k - k_{ss}) + 0.29\,(z^e - z^e_{ss}) + ... \tag{29}$$

$$l - l_{ss} = -0.002\,(k - k_{ss}) + 0.29\,(z^e - z^e_{ss}) + ... \tag{30}$$

$$k' - k_{ss} = 0.97\,(k - k_{ss}) + 0.67\,(z^e - z^e_{ss}) + ... \tag{31}$$

From these optimal decision rules, equations (29), (30), and (31), it appears that when agents are unsure about the the state of the economy the weight on z^e relative to k is lower compared to the complete information case for all the decision rules except hours. This directly reflects the impact of uncertainty on endogenous decisions.[16]

---

[16]Also note that as we move from the complete information case to the incomplete information case, while the coefficients on capital in these equations remain unchanged, the coefficients on technology change.

As in the case of complete information, the reason why the variability of the endogenous variables changes as we move from high to low volatility economies is primarily due to changes in the volatility of $(k - k_{ss})$ and $(z^e - z^e_{ss})$. However, unlike the complete information case, now the volatility of both $(k - k_{ss})$ and $(z^e - z^e_{ss})$ falls more than proportionately relative to the decline in the volatility of z as we move from high to low volatility economies—see the bottom panel in Table 3. Moreover, it is the relative volatility of z^e that declines sharply—the standard deviation of z^e relative to the unconditional standard deviation of z is only 14 percent $(0.30/2.15)$ in the low versus the high volatility economy.[17]

Our conjecture is that the sharp fall in the relative volatility of z^e is primarily due to a harder inference problem in the low volatility economy. The following subsection investigates this further.

Uncertainty and the expected state From equation (14), it is clear that the volatility of the expected level of technology, z^e, depends largely on the volatility of the expected state. In figures 2 and 3, we plot time series of the latent state s_t and the agent's expectations of that state at each date, s_t^e. The state s_t is either 0 or 1 and is indicated by solid diamonds at 0 and 1 in the figures. The expectation is indicated by the gray triangles and is never exactly zero or one. For the high volatility economy, shown in Figure 2, the agent is only rarely confused about the state. This is characterized by relatively few dates at which the expectation of s_t is not close to zero or one. Consequently, the expected level of the state and hence the expected level of technology are more volatile in this case. For the low volatility economy,

---

[17]Because the volatility of k does not fall as much as the volatility of z^e, the coefficients in the decision rules matter much more in the incomplete information case with Bayesian learning. In the labor hours rule, the ratio of coefficients between z^e and k is on the order of $150 : 1$. For consumption, it is about $10 : 1$, and for capital it is closer to $1 : 1$. Accordingly, the decline in z^e volatility relative to z volatility has a dramatic impact on hours volatility, a significant impact on consumption volatility, and a more moderate impact on capital volatility. Consequently, learning accounts for a large fraction, 62 percent, of the decline in hours volatility, a more modest fraction, 46 percent, of the decline in consumption volatility, and a smaller fraction, 25 percent, of the decline in investment volatility.
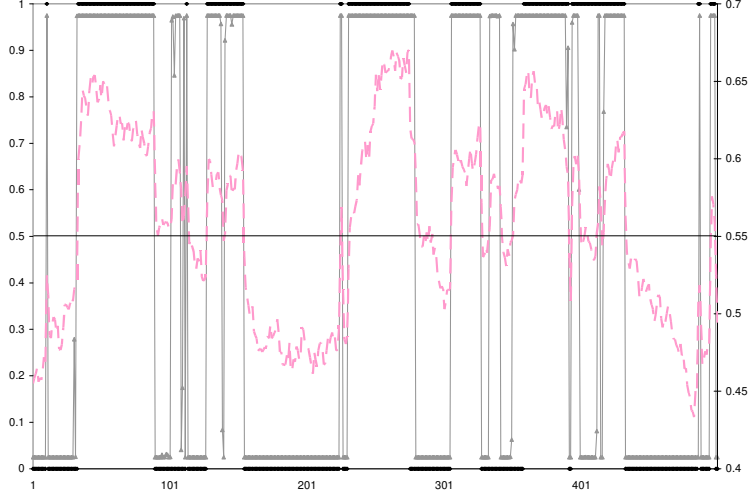
Figure 2: The true state s_t versus the expected state in the high volatility economy, measured on the left scale. The true state is indicated by solid diamonds at zero or one. The expected state is represented by the gray triangles. The dashed line shows the evolution of the log of output about its mean value of 0.55, measured on the right scale. The agent is relatively sure of the state in this economy.

shown in Figure 3, the agent is confused about the state much more often, as indicated by many more dates at which the expectation of the state is far from zero or one—more gray triangles nearer 0.5. This leads to a less volatile s_t^e and z_t^e.

These figures also show the evolution of output for each economy. The log of output is measured on the right scale in the figures and is shown as a dashed line. The logarithm of the steady state of output is 0.55 and is shown as a solid line; we can therefore refer to output above or below steady state. Output tends to be above steady state when beliefs are high and below steady state when beliefs are low.

A surprise Confusion about the latent state s_t leads to some surprising behavior which we did not expect to find. This behavior is illustrated in
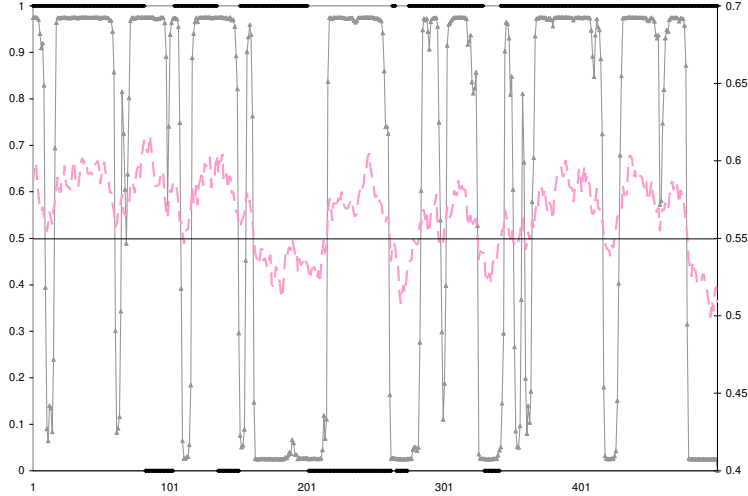
24

Figure 3: The true state versus the expected state in the low volatility economy, along with the evolution of log output about its steady state value. The agent is relatively confused about the true state, causing moderated behavior.

Figure 3. In particular, the agent sometimes believes in recession or expansion states when in fact the opposite is true. This occurs, for instance, in the time period around $t = 250$ in this simulation. Here the true state is low, but the agent believes the state is high. Interestingly, output remains above steady state for this entire period. The beliefs are driving the consumption, investment, and labor supply behavior of the agent in the economy, such that belief in the high regime is causing output to boom.[18]

How does this belief-driven behavior come about? At the end of each period, agents can observe labor, capital, and output and therefore can infer a value for z_t. Let's suppose the agent observes a high level of labor input and a high level of output. The agent may infer that the current la-

---

[18]We also calculated the real wage and interest rate volatility to see if prices adjust more than one-for-one in the low volatility period to compensate for the discrepancy between the actual and the expected state, and therefore the expected level of technology. We find that relative to the shock process, real wages are more volatile in the low volatility period, but real interest rates are less volatile.

25

tent state s_t is high and construct next period's expectation of the level of technology based on the expectation that s_{t+1} is also likely to be high (since the latent state is very persistent). But the high level of labor input may also itself have been due to an expectation of a high level of technology in the past period. The agent may therefore propagate the expectation of a high state forward. Labor input in the current period would then again be high, output would again be high, and the agent may again infer that the state s_t is high and construct next period's expectation of the level of technology based on the expectation that s_t is high. In this way beliefs can influence the equilibrium of the economy, and this effect is more pronounced as regimes move closer together.

Another way to gain intuition for the nature of the belief-driven behavior is to consider equation (18), which is derived earlier and reproduced here:

$$z^e_{t+1} = f\left(z^e_t, z_t\right). \tag{32}$$

The expected level of technology is a state variable in this system. The agent is able to calculate a value for z_t at the end of each period after production has occurred based on observed values of y_t, k_t, and ℓ_t, and this provides an input, but not the only input, into the next period's expected level of technology. This is because the decisions taken today that produced today's output depend in part on the belief that was in place at the beginning of the period, z^e_t. The true state is not fully revealed by the z_t calculated at the end of the period. Nevertheless, when regimes are far apart, the evidence is fairly clear regarding which state the economy is in and so z_t provides most of the information needed to form an accurate expectation z^e_{t+1}. When regimes are closer together, z_t is not nearly as informative and the previous expectation z^e_t can play a large role in shaping z^e_{t+1}.

4 An estimated shock process

4.1 Overview

Until this point, we have considered a calibrated case in which the stochastic driving process changes in such a way (by moving the regimes closer together) that equilibrium output volatility falls by 50 percent, as suggested by the data. Other aspects of the calibration were chosen to remain consistent with standards in the equilibrium business cycle literature. In this section we take an alternative approach. We estimate the stochastic driving process in a manner similar to Kim and Nelson (1999a),[19] and then examine the implied volatility reduction and the component of that reduction that can be attributed to the learning effect.

4.2 Data

Table 4 reports the business cycle volatility of key variables that are relevant for our analysis.[20] In this table we also compare how volatile the 1984:1-2004:4 period is relative to the 1954:1-1983:4 period. Based on the last column we note that (i) overall all the variables are less volatile after 1984, and (ii) the data suggests that reduction in business cycle volatility is not equal across different macroeconomic variables. Any satisfactory explanation for this reduction in volatility must then endogenously produce an asymmetric response for different series. A standard real business cycle model would not be compelling in this respect for reasons mentioned in our discussion of Figure 1. Below we report the asymmetric effects in our model, some of which are promising, and others of which will call for further additions to the model to match data.

---

[19]We depart from Kim and Nelson (1999a) in two ways. We fit the regime-switching process to Hodrick-Prescott filtered technology and we assume an exogenous structural break.

[20]We use quarterly data and the sample period is 1954:1-2004:4. All National Income and Product Accounts (NIPA) data from the Bureau of Economic Analysis (BEA) is in billions of chained 2000 dollars. The employment data is from the establishment survey. Data has been logged before being detrended using the Hodrick-Prescott filter.

TABLE 4. CYCLICAL VOLATILITY OF KEY VARIABLES

	1954:1-83:4	1984:1-04:4	Low/High
GNP	1.92	0.95	0.50
Personal consumption expenditure	0.92	0.73	0.79
Nondurables and services consumption	0.91	0.57	0.63
Consumption of durables	4.91	2.94	0.60
Gross private investment	8.39	5.16	0.62
Fixed investment plus consumer durables	5.34	3.14	0.59
Total hours of work	1.80	1.08	0.60
Average weekly hours of work	0.50	0.37	0.74
Employment	1.57	0.94	0.60
TFP	1.80	1.02	0.57

Table 4: Percent standard deviation of the cyclical component of key macro-economic variables.

4.3 Estimates of the technology process

The total factor productivity (TFP) series is constructed using $log(z_t) = \log(GNP_t) - (1 - \alpha)\log(Hours_t) - \alpha\log(Capital_t)$.[21,22] We then fit a regime-switching process on total factor productivity after detrending it using the Hodrick-Prescott filter for the sample period 1954:1 to 2004:4 with 1983:4 as an exogenous break date. To stay consistent with Kim and Nelson (1999a) we hold the conditional (within regime) standard deviation and the transition probabilities of the Markov switching process constant across the two sample periods. We also assume that the process is symmetric.[23] An advan-

---

[21]The measure of output used here is real GNP which is in chained 2000 dollars. The labor input is measured in aggregate hours. We construct the series for aggregate hours by multiplying payroll employment data with average weekly hours. As a measure of capital, we use the gross stock of real nonresidential fixed private capital.

[22]To stay consistent with the business cycle model used here, this measure does not adjust for cyclical variations in labor effort and capacity utilization. Such abstractions, as has been previously noted in the literature (see King and Rebelo (1999) for an earlier overview of the main drawbacks of using the Solow residual as a measure of aggregate technology) can lead to significant mis-measurements in the measure of technology.

[23]Results for a model with an asymmetric regime-switching process are reported in Table 8 and Table 9 in the Appendix. We find that the asymmetric case does not fit the data significantly better than the symmetric case because the likelihood ratio statistic is 0.348, substantially less than the critical value of 7.81 at the 95 percent significance level. Several empirical studies document asymmetries in business cycles, both in terms of duration and

TABLE 5. ESTIMATES OF THE TECHNOLOGY PROCESS.

	1954:1-1983:4	1984:1-2004:4
	Parameter Values	
Regime distance	2.709 (0.195)	0.657 (0.389)
Transition probability	0.844 (0.034)	
Conditional standard deviation	1.116 (0.129)	
Log likelihood	-339.69	

Table 5: Estimates of the coefficients of the regime-switching process for the two sample periods with an exogenous break date of 1983:4. Standard errors are in parenthesis. We allow only regime distance to change across the two samples. Transition probability and conditional standard deviation are held constant across the two sample periods. The regime distance and conditional standard deviation are expressed as percent.

tage of this approach is that the business cycle volatility generated by the estimated technology process can be compared to the calibrated economies studied in the previous section.

We estimate four parameters: Regime distance for both sample periods, transition probability and conditional standard deviation. Table 5 reports our estimates.

4.4 Moderation in the baseline estimated case

Using the estimates of our technology process we compute a_H, a_L, and the parameters of the stochastic $AR(1)$ process for z_t given by equation (10). The top panel of Table 6 reports these parameters. The high and low states of technology have come closer together after 1984, as reflected in the top panel. However, the unconditional standard deviation of the technology process declines by 20 percent instead of the 35 percent decline in our calibrated example.[24] Combining the estimates of the process of technology

---

amplitude. However, after filtering the data using Hodrick-Prescott filter typically there is very little evidence of asymmetry.

[24]The cyclical measure of TFP was half as volatile after 1984 as shown in the last row in Table 4. However, our estimation procedure captures about 45 percent of the actual decline

TABLE 6. MODERATION IN THE ESTIMATED CASE.

	High volatility economy	Low volatility economy
	Parameter Values	
$a_H = a_L$	0.014	0.003
$p = q$	0.844	0.844
$\sigma\sigma_\epsilon$	0.0162	0.013
$\sigma\varsigma$	0.011	0.011
	Volatility, in percent standard deviation	
Output	2.044	1.335
Consumption	0.289	0.147
Hours	0.616	0.071
Investment	7.716	5.166

Table 6: Comparison of the business cycle volatility in the high volatility and low volatility incomplete information economies. The top panel reports the relevant parameters of the technology process based on our estimation. The bottom panel reports the percent standard deviation of the Hodrick-Precott filtered data of key endogenous variables.

with the calibrated values of the rest of the parameters we compute the cyclical volatility implied by our model.

According to these estimates, the unconditional standard deviation of the technology shock fell by 20 percent across the two periods $(0.013/0.0162)$. As a consequence, the volatility of output fell, albeit by 35 percent $(1.335/2.044)$. This 35 percent reduction in output volatility is about two-thirds of the actual moderation observed in the data. Still, of the estimated moderation in output volatility of 35 percent, a significant component, about 43 percent,

---

observed in the data. The key reason for this is that we restrict the conditional standard deviation to remain constant across the two sample periods to stay consistent with Kim and Nelson (1999a). If we remove this restriction, the variability of the TFP process declines by 50 percent, the model collapses to a linear model—there is no role for learning—and the variability of output declines by 50 percent as well. Such an outcome is not surprising and is consistent with the findings of Arias, Hansen and Ohanian (2007).

However, as noted earlier, we want to take the core finding of Kim and Nelson (1999a) as a primitive for our quantitative-theoretic analysis. Our analysis explores the role of learning when there is imperfect information in an otherwise standard equilibrium business cycle model. Imperfect information with Bayesian learning allows us to move away from the case where there is one-to-one correspondence between the variability of the shock and the variability of the endogenous variables.

TABLE 7. SERIAL CORRELATIONS.

	Baseline estimated case		Data	
	High volatility	Low volatility	High volatility	Low volatility
Output	0.62	0.54	0.84	0.87
Consumption	0.79	0.90	0.73	0.80
Hours	0.58	0.81	0.90	0.93
Investment	0.59	0.53	0.90	0.95

Table 7: Comparing serial correlation of the cyclical component of key variables. In the data, the serial correlation of output, consumption, hours and investment corresponds to the serial correlation of real GNP, consumption expenditure on non-durables and services, total hours in the establishment survey, and fixed investment plus durable goods consumption.

is due to learning $(1 - 20/35)$.

The moderations produced for other variables differ from those for output. Learning produces asymmetric effects across different macroeconomic variables, some of which are almost in line with what we observe in the data.[25] For consumption, the reduction of 49 percent exceeds the volatility reduction observed in the data as described in Table 4. However, with respect to investment, the model generated volatility decline is close to what we observe in the data. For hours the decline is almost 88 percent, far exceeding what we observe in the data.

4.5 Change in correlations

In our analysis so far we have focused on how incomplete information impacts business cycle volatility. We now turn to examine the implications for serial correlation for the baseline estimated case.[26] In Table 7, we report the first-order serial correlations for the cyclical component of key macro-

---

[25]To stay consistent with the model and the literature, we compare consumption, hours and investment generated by the model to their respective components in the data: consumption of non-durables and services, total hours input from establishment survey, and fixed investment plus durable goods consumption.

[26]Van Nieuwerburgh and Veldkamp (2006) interpret their related model as capturing correlations of expected variables on the grounds that actual variables are not observed contemporaneously in actual economies. We have not pursued this approach here.

economic variables for the estimated baseline case and compare them to the data. The serial correlations implied by the model are lower than in a standard equilibrium business cycle model. This is not surprising as the $AR(1)$ coefficient for the stochastic technology process is 0.72 in our baseline estimated case whereas this coefficient is 0.95 in a standard equilibrium business cycle model. In the data, the serial correlation in the low volatility period has increased for all the variables considered here. The incomplete information model generates an increase in serial correlation in the low volatility period for consumption and hours, but not for output and investment.[27]

5   Conclusions

We have investigated the idea that learning may have contributed to the great moderation in a stylized regime-switching economy. The main point is that direct econometric estimates may overstate the degree of "luck" or moderation in the shock processes driving the economy. This is because the changes in the nature of the shock process with incomplete information can also change private sector behavior and hence the nature of the equilibrium. Our complete information model has provided a benchmark in which it is well known that equilibrium volatility is linear in the volatility of the shock process, such that doubling the volatility of the shock process will double the equilibrium volatility of the endogenous variables. Against this background, we have demonstrated that learning introduced a pronounced nonlinear effect on volatility, in which private sector behavior changes markedly in response to a changed stochastic driving process for the economy with incomplete information. We have found, in a benchmark calculation, that such an effect can account for about 30 percent of a change in observed volatility. We think this is substantial and is worth in-

---

[27]We also considered the implications for contemporaneous correlations of macroeconomic variables with output. In moving from high to low volatility economies, the contemporaneous correlation for consumption changes from 0.3 to 0.0, for hours 0.6 to 0.5, and for investment 1.0 to 1.0.

vestigating in more elaborate models that can confront the data along more dimensions.

References

[1] Ahmed, S., A. Levin and B. Wilson. 2004. "Recent U.S. Macroeconomic Stability: Good Luck, Good Policies, or Good Practices?" *Review of Economics and Statistics* 86(3): 824-32.

[2] Andolfatto, D. and P. Gomme. 2003. "Monetary Policy Regimes and Beliefs." *International Economic Review* 44: 1-30.

[3] Arias, A., G. Hansen, and L. Ohanian. 2007. "Why Have Business Cycle Fluctuations Become Less Volatile?" *Economic Theory* 32(1): 43-58.

[4] Aruoba, S.B., J. Fernandez-Villaverde, and J.F. Rubio-Ramirez. 2006. "Comparing Solution Methods for Dynamic Equilibrium Economies." *Journal of Economic Dynamics and Control*, 30(12): 2477-2508.

[5] Burns, A. 1960. "Progress Towards Economic Stability." *American Economic Review* 50(1): 1-19.

[6] Cagetti, M., L. Hansen, T. Sargent and N. Williams. 2002. "Robustness and Pricing with Uncertain Growth." *Review of Financial Studies* 15(2): 363-404.

[7] Campbell, S. 2007. "Macroeconomic Volatility, Predictability, and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters." *Journal of Business and Economic Statistics* 25(2): 191-200.

[8] 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.

[9] Cooley, T.F., and E.C. Prescott (1995). "Economic Growth and Business Cycles." In Cooley T. F. (ed) *Frontiers of Business Cycle Research*. Princeton University Press, 1-38.

[10] Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

[11] Fernandez-Villaverde, J., and J. Rubio-Ramirez. 2007. "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies*, 74(4): 1059-1087.

[12] Hamilton, J. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica* 57(2): 357-384.

[13] Justiniano, A., and G. Primiceri. 2008. "The Time-Varying Volatility of Macroeconomic Fluctuations." *American Economic Review*, 98(3): 604-641.

[14] Kahn, J., M. McConnell, and G. Perez-Quiros. 2002. "The Reduced Volatility of the U.S. Economy: Policy or Progress?" Manuscript, FRB-New York.

[15] Kim, C.-J., and C. Nelson. 1999a. "Has the U.S. Economy become more Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *The Review of Economics and Statistics* 81(4): 608-616.

[16] Kim, C.-J., and C. Nelson. 1999b. *State Space Models with Regime Switching*. MIT Press, Cambridge, MA.

[17] Kim, C.-J., C. Nelson, and J. Piger. 2004. "The Less-Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations." *Journal of Business and Economic Statistics* 22(1): 80-93.

[18] King, R. and S. Rebelo (1999). "Resuscitating Real Business Cycles." *Handbook of Macroeconomics*. Volume 1B. North-Holland: Amsterdam. pp. 927-1007.

[19] McConnell, 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.

[20] Milani, F. 2007. "Learning and Time-Varying Macroeconomic Volatility." Manuscript, UC-Irvine.

[21] Moore, G.H. and V. Zarnowitz. 1986. "The Development and Role of the NBER's Business Cycle Chronologies." In: Robert J. Gordon, ed., *The American Business Cycle: Continuity and Change*. Chicago: University of Chicago Press.

[22] Owyang, M., J. Piger, and H. Wall. 2008. "A State-Level Analysis of the Great Moderation." *Regional Science and Urban Economics*, 38(6): 578-89.

[23] Primiceri, G. 2005. "Time-Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies* 72(3): 821-852.

[24] Van Nieuwerburgh, S. and L. Veldkamp. 2006. "Learning Asymmetries in Real Business Cycles." *Journal of Monetary Economics* 53(4): 753-772.

[25] Sims, C., and T. Zha. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* 96(1): 54-81.

[26] Stock, J., and M. Watson. 2003. "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual 2002* 17: 159-218.

A Beliefs and expectations

A.0.1 Beliefs

We follow Kim and Nelson (1999b) in the following discussion of the evolution of beliefs. At date t, agents forecast s_{t-1} given information available at date t. Letting $b_t = P(s_{t-1} = 0|F_t)$,

$$b_t = \sum_{s_{t-2}=0,1} P(s_{t-1} = 0, s_{t-2}|F_t)$$
$$= P(s_{t-1} = 0, s_{t-2} = 0|F_t) + P(s_{t-1} = 0, s_{t-2} = 1|F_t), \quad (33)$$

where the joint probability that the economy was in a recession in the last two periods is given by

$$P(s_{t-1} = 0, s_{t-2} = 0|F_t) = P(s_{t-1} = 0, s_{t-2} = 0|z_{t-1}, F_{t-1})$$
$$= \frac{\phi(z_{t-1}, s_{t-1} = 0, s_{t-2} = 0|F_{t-1})}{\phi(z_{t-1}|F_{t-1})}$$
$$= \frac{\phi_L(z_{t-1}|s_{t-1} = 0, s_{t-2} = 0, F_{t-1})}{\phi(z_{t-1}|F_{t-1})} \times$$
$$P(s_{t-1} = 0|s_{t-2} = 0, F_{t-1})P(s_{t-2} = 0|F_{t-1}), \quad (34)$$

where ϕ_i denotes the density function under regime $i \in \{L, H\}$, and $\phi(z_{t-1} | F_{t-1}) = \Sigma_{s_{t-1}}\Sigma_{s_{t-2}}\phi(z_{t-1}, s_{t-1}, s_{t-2}|F_{t-1})$. Similarly,

$$P(s_{t-1} = 0, s_{t-2} = 1|F_t) =$$
$$\frac{\phi_L(z_{t-1}|s_{t-1} = 0, s_{t-2} = 1, F_{t-1})P(s_{t-1} = 0|s_{t-2} = 1, F_{t-1})P(s_{t-2} = 1|F_{t-1})}{\phi(z_{t-1}|F_{t-1})}.$$
$$(35)$$

Using the transition probabilities define g_L and g_H as

$$g_L = \phi_L(z_{t-1}|s_{t-1} = 0, s_{t-2} = 0, F_{t-1})qb_{t-1}$$
$$+ \phi_L(z_{t-1}|s_{t-1} = 0, s_{t-2} = 1, F_{t-1})(1 - p)(1 - b_{t-1}), \quad (36)$$

and

$$g_H = \phi_H(z_{t-1}|s_{t-1} = 1, s_{t-2} = 0, F_{t-1})(1 - q)b_{t-1}$$
$$+ \phi_H(z_{t-1}|s_{t-1} = 1, s_{t-2} = 1, F_{t-1})p(1 - b_{t-1}). \quad (37)$$

Since $z_{t-1} = \xi_0 + \xi_1 z_{t-2} + \sigma(v_{t-1} + \varsigma \eta_{t-1} - \lambda_1 \varsigma \eta_{t-2})$, then conditional on v_t, z_t has a normal distribution. Letting $s_{t-1} = 0$ and $s_{t-2} = 0$, then $v_{t-1} = -(1-q)$ and $z_{t-2} = -a_L + \sigma \varsigma \eta_{t-2}$, and so, if in the last two periods the economy was in a recession, $z_{t-1} = \xi_0 + \xi_1(-a_L) - \sigma(1-q) + \sigma \varsigma \eta_{t-1}$. We can therefore write the conditional density function as

$$\phi_{L00} = \phi_L(z_{t-1} | s_{t-1} = 0, s_{t-2} = 0, F_{t-1})$$
$$= \frac{1}{\sqrt{2\pi\sigma^2\varsigma^2}} \exp\left(-\frac{(z_{t-1} - \xi_0 - \xi_1(-a_L) + \sigma(1-q))^2}{2\sigma^2\varsigma^2} \right). \quad (38)$$

When $s_{t-1} = 0$ and $s_{t-2} = 1$, then $v_{t-1} = -p$ and $z_{t-2} = a_H + \sigma \varsigma \eta_{t-2}$, and the density function is

$$\phi_{L10} = \phi_L(z_{t-1} | s_{t-1} = 0, s_{t-2} = 1, F_{t-1})$$
$$= \frac{1}{\sqrt{2\pi\sigma^2\varsigma^2}} \exp\left(-\frac{(z_{t-1} - \xi_0 - \xi_1(a_H) + \sigma p)^2}{2\sigma^2\varsigma^2} \right). \quad (39)$$

Similarly

$$\phi_{H01} = \phi_H(z_{t-1} | s_{t-1} = 1, s_{t-2} = 0, F_{t-1})$$
$$= \frac{1}{\sqrt{2\pi\sigma^2\varsigma^2}} \exp\left(-\frac{(z_{t-1} - \xi_0 - \xi_1(-a_L) - \sigma q)^2}{2\sigma^2\varsigma^2} \right), \quad (40)$$

and

$$\phi_{H11} = \phi_H(z_{t-1} | s_{t-1} = 1, s_{t-2} = 1, F_{t-1})$$
$$= \frac{1}{\sqrt{2\pi\sigma^2\varsigma^2}} \exp\left(-\frac{(z_{t-1} - \xi_0 - \xi_1(a_H) - \sigma(1-p))^2}{2\sigma^2\varsigma^2} \right). \quad (41)$$

Thus we can write b_t as

$$b_t = \frac{g_L}{g_L + g_H}. \quad (42)$$

A.0.2 Expectations

Since b_t is the probability that the economy was in a recession and $(1 - b_t)$ is the probability that the economy was in an expansion in the last period,

TABLE 8. ESTIMATES OF THE TECHNOLOGY PROCESS FOR ASYMMETRIC CASE.

	1954:1-1983:4	1984:1-2004:4
Parameter Values		
$a_H + a_L$	2.713 (0.195)	0.703 (0.448)
a_L	-1.406 (0.212)	-0.276 (0.214)
p		0.848 (0.050)
q		0.844 (0.052)
σ_ς		1.108 (0.132)
Log likelihood		-339.52

Table 8: Estimates of the coefficients of the regime-switching process for the asymmetric case. Standard errors are in parenthesis. The regime distance and the conditional standard deviation are expressed as percent.

we determine the probability distribution of the current state by allowing for the possibility of state change. In particular,

$$[P(s_t = 0|F_t), P(s_t = 1|F_t)]$$

$$= [P(s_{t-1} = 0|F_t), P(s_{t-1} = 1|F_t)] \begin{bmatrix} q & 1-q \\ 1-p & p \end{bmatrix} \quad (43)$$

which can be rewritten as

$$[P(s_t = 0|F_t), P(s_t = 1|F_t)] = [b_t, (1-b_t)] \begin{bmatrix} q & 1-q \\ 1-p & p \end{bmatrix}. \quad (44)$$

Given that $P(s_t = 0|F_t) = b_t q + (1 - b_t)(1 - p)$ and $P(s_t = 1|F_t) = b_t(1 - q) + (1 - b_t)p$,

B Estimates of the asymmetric technology process

In the asymmetric case we no longer impose that $a_H = a_L$ and $p = q$. As before, we hold the transition probabilities and conditional standard deviation constant across the two sample periods. Therefore, we estimate 7 parameters here: $a_H + a_L$, a_L for each sample period, transition probabilities p, q, and conditional standard deviation. Table 8 reports our estimates.

TABLE 9. MODERATION IN THE ASYMMETRIC MODEL.

	High volatility economy	Low volatility economy
	Parameter values	
a_H	0.013	0.004
a_L	0.014	0.003
p	0.848	0.848
q	0.844	0.844
$\sigma\sigma_\epsilon$	0.016	0.013
σ_ς	0.011	0.010
	Volatility, in percent standard deviation	
Output	1.960	1.479
Consumption	0.276	0.164
Hours	0.603	0.089
Investment	7.408	5.715

Table 9: Comparison of the business cycle volatility in the high volatility and low volatility incomplete information economies. The top panel reports the relevant parameters of the technology process based on our estimation of the asymmetric case. The bottom panel reports the percent standard deviation of the Hodrick-Precott filtered data of key endogenous variables.

Using these estimates, we compute the remaining parameters of the technology process, reported in the top panel of Table 9 and the bottom panel reports the business cyclical volatility implied by these estimates.