

NBER WORKING PAPER SERIES

STRUCTURAL BREAKS IN AN ENDOGENOUS GROWTH MODEL

Timothy Cogley
Boyan Jovanovic

Working Paper 28026
<http://www.nber.org/papers/w28026>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020

We thank the NSF and the C.V. Starr Center for support, F. Bianchi, R. Lucas, S. Ludvigson, V. Midrigan, P. Rousseau, L. Veldkamp and G. Violante for comments, and Gaston Navarro, Sai Ma, and Angelo Orane for assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Timothy Cogley and Boyan Jovanovic. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Structural Breaks in an Endogenous Growth Model
Timothy Cogley and Boyan Jovanovic
NBER Working Paper No. 28026
October 2020
JEL No. E1

ABSTRACT

We study the effects of parameter uncertainty prompted by structural breaks. In our model, agents respond differently to uncertainty prompted by regime shifts in shock processes than they react to comparable perceived increases in shock volatility. The magnitude of the response to an increase in uncertainty about TFP associated with a structural break is greater than that of a response to a comparable perceived rise in volatility. This is because lifetime utility varies more when shocks shift beliefs and perceived wealth.

Timothy Cogley
Department of Economics
New York University
19 West Fourth Street, Sixth Floor
New York, NY 10012
and NBER
tim.cogley@nyu.edu

Boyan Jovanovic
New York University
Department of Economics
19 W. 4th Street, 6th Floor
New York, NY 10012
and NBER
Boyan.Jovanovic@nyu.edu

1 Introduction

Are learning episodes similar to times when shock volatilities rise? We study this question in a general-equilibrium model for an economy that is subject to recurrent structural breaks, breaks that create parameter uncertainty. Our model is based on an Ak growth model with two aggregate shocks, a neutral TFP shock in the final-goods sector and an investment-specific technology shock that affects production of capital goods. We deal only with aggregate shocks and abstract from idiosyncratic risk.

Structural breaks are introduced as occasional shifts in the parameters governing technology shocks. A break alters medium- and long-run forecasts of the level of technology and initiates a period of higher parameter uncertainty. We examine how this uncertainty affects savings, investment, and growth. Under uncertainty, beliefs over the distributions of future shocks are an aggregate state that responds to current shock realizations. Shocks thus have wealth effects through the beliefs channel – a channel that is absent in models where the change in the shock distributions is known, and one that creates both a quantitative and a qualitative difference between the two classes of models.

To motivate interest in structural breaks, figure 1 plots empirical counterparts of the shock processes¹ along with *NBER* recession dates, shown as vertically shaded areas. Ten-year rolling standard deviations are depicted in the bottom row.

TFP is shown in the upper left panel. As highlighted by Gordon (2016, pp. 545-548), productivity surges in the 1940s and never returns to its pre-war level. Although TFP also varies procyclically, the mid-century surge is the dominant source of volatility and presumably also of uncertainty. For instance, the rolling 10-year standard deviation is about 5 times higher around the years of the rise than at ‘normal’ times.²

¹In our model, TFP equals the ratio of real output to real capital. Appendix A describes our data sources for y/k .

Data for aggregate q come from Wright (2004). His “equity q ” measure covers the period 1900-2002, and we extend this through 2018 by ratio splicing the Federal Reserve Board’s measure of equity q (Financial Accounts of the United States - Z.1, table B.103, line 45).

²The same pattern holds if multifactor TFP is used instead of the output-capital ratio: Fig. 16-5 of Gordon (2016) shows that the growth of TFP over the 1940-50 decade was about twice as high as it was over the adjacent decades, and about four times as high as it was over the other decades since 1900.

Data in Piketty (2014) suggest that structural breaks in Y/K also occurred in Britain and France, although perhaps at different times (compare Piketty’s tables 3.1 and 3.2 for Britain and France

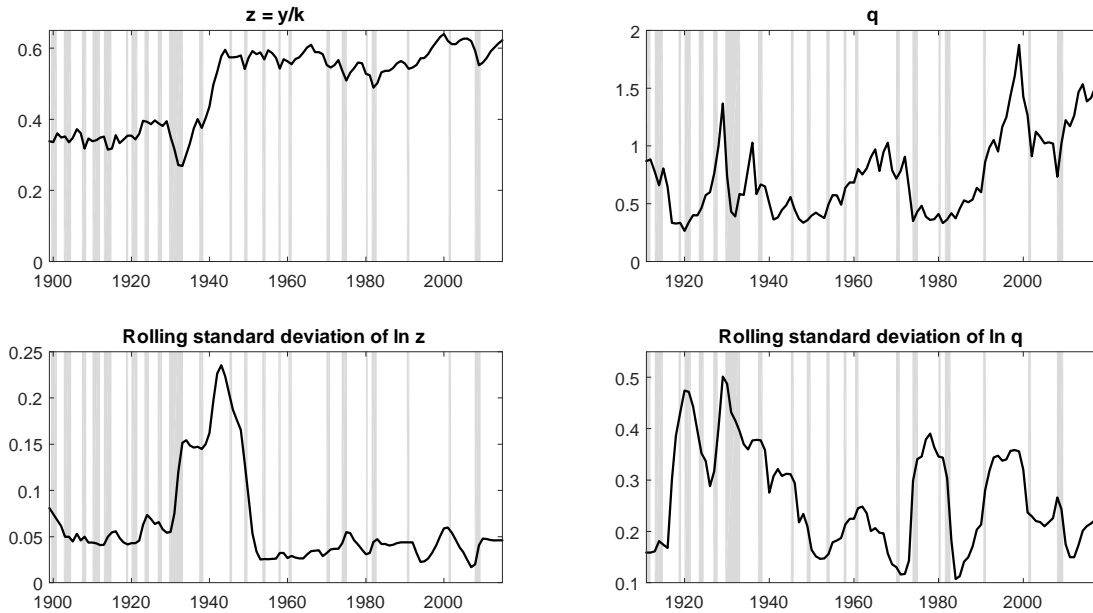


Figure 1: TFP, z , AND INVESTMENT COST, q , OVER A CENTURY. LEVELS ARE DEPICTED IN THE TOP ROW, AND 10 YEAR ROLLING STANDARD DEVIATIONS ARE SHOWN IN THE BOTTOM. SHADED AREAS MARK THE DATES OF NBER RECESSIONS.

The case for a structural break in the investment-specific technology shock – shown in the upper right panel – is less obvious, but an argument could be made that the run-up in the 1990s and early 2000s represented a shift in its mean. For instance, the peak and trough after 1995 are two and three times higher, respectively, than those of the previous 90 years, and the trough in 2002 is almost as high as the peaks in the 1920s and 1960s. The volatility associated with this shift is comparable to that experienced in earlier years, neither dominating nor being dominated by ordinary cyclical variation.³

Our goal is to understand how shifts such as these affect savings, investment, and growth in general equilibrium. Our analysis is based on three assumptions. The first is that if a structural break has occurred in the past, others are possible in the future. Furthermore, agents are aware of this possibility and account for it when formulating their plans. Structural shifts are not complete surprises and as Sims (1982) argued

with his table 4.6 for the US).

³We acknowledge that this is not the only respectable interpretation of the data. Alas, because statistical tests for structural breaks have low power, a decisive characterization is unlikely.

in a related context, regime shifts can be viewed as outcomes of a rule that agents understand. This contrasts with anticipated-utility models in which agents update beliefs after a structural break but fail to account for the possibility of future breaks when formulating plans (e.g., Kreps 1998 or Evans and Honkapohja 2001).

Second, history never repeats itself exactly. Regime realizations do not necessarily cycle between a small number of possible states, as in the literature on Markov-switching models. Entirely new regimes can and do emerge. The timing of the shift follows a Poisson process and conditional on a structural break new parameter values are drawn from a distribution conditioned on the previous parameter realizations; the exact nature of regimes is first-order Markov process on the space of shock-distribution parameters.

Third, because each realization is distinct, a structural break initiates a period of uncertainty. The new regime is not suddenly revealed after a break. On the contrary, agents must learn the new parameters governing the shocks. It follows that a structural break also raises uncertainty. A structural break combines a persistent change in the level of technology with a transitory increase in uncertainty. Both affect behavior and macroeconomic outcomes. Using our model, we can quantify and disentangle them.

The model has no heterogeneity among agents, and our study is therefore limited to the effects of macro uncertainty and macro risk. The model also has no financial sector and would be inappropriate for analyzing the great recession, and our sample ends in 2006. On the other hand, the model is solved in closed form, without any linearization. Moreover the Ak structure allows a direct mapping between the observables and the shocks, and the endogeneity of growth amplifies the transitional dynamics will be different – precautionary savings will, for instance, have a stronger effect on future output.

For plausible calibrations, we find that uncertainty episodes entail lower consumption and higher growth compared the economy that is identical except that agents know the parameters of the distributions governing the shocks. Uncertainty prompts precautionary savings that insures agents against an unfavorable change. Consumption falls and the higher savings then raise growth above its no-uncertainty level. Each type of learning, be it about TFP or about investment efficiency, has an expansionary effect on growth, although the effect is smaller in the second case. Uncertainty about TFP leads to more saving because in our closed economy low-TFP realizations lead

to consumption disasters whereas high-investment-cost realizations do not because agents can reduce investment accordingly.⁴

To isolate the effects of a rise in uncertainty, however, we then ask whether these effects arise because agents live through episodes during which they are learning the parameters of shock distributions, or whether they would occur even if agents simply faced higher shock variances. The effect on growth of a mean-preserving increase in variance (MPS) that agents immediately understand can go in either direction. An MPS activates two forces: higher investment as a consequence of precautionary savings versus lower investment because of the option to wait (e.g., Dixit and Pindyck 1994). Shock persistence matters, however, because waiting is more attractive when shocks are weakly persistent. Since Bayesian learning makes one of the state variables a martingale (Doob 1948), a strongly persistent state is present in all our learning models. The presence of a persistent state variable weakens incentives to wait and brings the precautionary-savings motive to the fore.⁵

The next step is to compare the magnitude of responses associated with subjective and objective uncertainty. Toward that end, we equate the sequences of one-step-ahead marginal distributions of shocks in the two cases. Under learning, subjective uncertainty falls over time, and the no-learning economy is put on the same footing by assuming a corresponding decline in the conditional variance of shocks. For plausibly calibrated models, consumption and growth respond more strongly under learning: after ten years, output is permanently higher by two percent. In contrast, a mean-preserving increase in variance matching the profile of subjective uncertainty has very little effect on savings or growth.

Models featuring TFP and investment-efficiency shocks but without learning include Greenwood, Hercowitz, and Krusell (2000), Fisher (2006), Justiniano, Primiceri and Tambalotti (2010) and Jovanovic and Rousseau (2014). Gilchrist and Williams (2005) find that a rise in dispersion of shocks over vintages of capital is expansionary because labor reallocates to the more productive vintages, but their MPS is cross-sectional. Bloom (2014) summarizes results on time-varying risk when shifts in the

⁴Even a bad TFP realization would not cause a consumption disaster in an open economy because agents can then borrow abroad. Baxter and Jermann (1997) note, however, that investors do not diversify internationally to any significant extent.

⁵Collin-Dufresne, et al. (2016) also emphasize the importance of this martingale state. They study implications for asset pricing in economies in which consumption growth is exogenous, whereas we study production economies in which uncertainty influences consumption and investment.

distributions of shocks are known.

Endogenous growth models have studied how growth responds to volatility in aggregate TFP shocks (Jones, Manuelli, Siu and Stacchetti 2005) and to policy shocks (Hopenhayn and Muniagurria 1996). We find that the qualitative effects of shocks to the efficiency of investment are similar to those of investment taxes or subsidies.

Among related models with learning, Huffman and Kiefer (1994) and Koulovantianos, Mirman and Santugini (2009) studied learning the TFP process in the production of final goods, but not that of capital goods. Bernanke (1983) and Stokey (2015) study partial equilibrium models with delayed information arrival about investment profitability. As in our model, learning occurs exogenously, as a result of the passage of time alone. It does not depend on actions that agents take – on how much to invest, for instance. This is in contrast to Chamley and Gale (1994), Veldkamp (2005), and Pastor and Veronesi (2009) whose models imply, as ours does, that the effects of learning differ from a rise in shock volatility.

Parameter learning also implies momentum effects on stock prices – serially correlated price changes caused by learning. As Lewellen and Schanken (2002) have shown, parameter uncertainty drives a wedge between the distribution perceived by investors and the distribution estimated by empirical tests. Following a regime shift, an econometrician should expect to find predictability of returns. As agents learn the parameters stock prices in general start to reflect this. For instance, if a regime shift lowered the distribution of TFP, returns would initially be low, reflecting the fall in dividends and would then gradually rise as stock prices fall to reflect the lower mean of the dividend process.

In our model momentum effects arise for the risk-free rate, but not for stock prices. If investment is positive, Tobin's Q (the market value of capital relative to its replacement cost) is unity. Only when investment hits zero can Q fall below unity. As Sargent (1980, p. 111) puts it, roughly speaking, Q drops farther below unity as the constraint that investment be irreversible becomes more binding.

In that case parameter uncertainty also leads to a temporary rise in the volatility of the stock market, and some regime changes may raise the possibility of disasters as in Veronesi (2004), Gourio (2008) and Wachter (2013). Regime shifts also raise uncertainty and if we had ambiguity-averse agents regime shifts would depress stock prices and raise returns that would gradually decline as learning takes place, as in Epstein and Schneider (2008).

The paper that is closest to ours is Bianchi and Melosi (2016). They study a family of DSGE models with Markov switching and Bayesian learning about latent states. Importantly, unlike anticipated-utility models of learning (Kreps 1998, Evans and Honkapohja 2001), their agents take regime shifts and learning fully into account when forming plans. For a real-business-cycle model in which TFP growth switches between high or low values, Bianchi and Melosi demonstrate that the anticipated-utility assumption is not innocuous because past mistakes about TFP growth are built into the capital stock and thus have persistent effects on output, consumption, and investment.⁶ We extend their analysis by breaking free of the assumption that there are finite number of potential states. In our model, there is a continuum of possible future states, and entirely new regimes can emerge. Because there are no best or worst possible states and period utility is unbounded, the potential influence of uncertainty is magnified.⁷

Last but not least, what we call “parameter uncertainty” differs from Knightian uncertainty because agents can still form predictive distributions over variables – they are Bayesians. Knight’s definition would more appropriately apply to boundedly rational agents. For example, Pintus and Suda (2018) feature adaptive learning and find, as we do for Bayes learning, that the responses of output, investment, and other aggregates under adaptive learning are significantly larger than when the agents know the parameters governing the shocks.⁸ One could also analyze the effects of shifts when agents have ambiguous beliefs (Hansen and Sargent 2001; Bianchi, Ilut, and Schneider 2018).

The paper is organized as follows. Section 2 describes an Ak growth model with these features and characterizes decision rules and the equilibrium law of motion. Section 3 presents a number of simple examples to build intuition. Section 4 solves and simulates calibrated versions in order to assess the strength of various forces, especially those governing uncertainty and precautionary behavior.

⁶Bianchi and Melosi (2018) use similar methods to examine the role of policy uncertainty in a new Keynesian model with recurrent shifts in the monetary-policy rule.

⁷For endowment economies, Geweke (2001) and Cogley (2009) examine the consequences for expected utility and the market price of risk, respectively.

⁸However, their adaptive-learning framework shuts down precautionary effects associated with uncertainty.

2 Model and Planner's Solution

Technology.—The model has two sectors: a final goods sector and a capital-goods sector. The production function for the final good is

$$Y = zk, \quad (1)$$

where Y is the output of final goods, k the capital-good input, and z a shock to the final goods technology. Capital depreciates at the rate δ ; the law of motion for capital is

$$k' = (1 - \delta)k + \frac{1}{q}X, \quad (2)$$

and the aggregate resource constraint is

$$Zk = C + X, \quad (3)$$

where X is final goods devoted to the production of capital, and q is the shock to the capital-goods technology.

Preferences.—A representative agent has recursive preferences as in Epstein and Zin (1989) and Weil (1989),

$$U_t = \left[(1 - \beta)C_t^{1-\rho} + \beta \left((E_t[U_{t+1}^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}} \right) \right]^{\frac{1}{1-\rho}}, \quad (4)$$

where β is the subjective discount factor, γ is the coefficient of relative risk aversion, and ρ is the inverse of the elasticity of intertemporal substitution.⁹ Expectations are taken with respect to subjective beliefs.

Aggregate resource constraint.—The Ak production structure in Eqs. (1) and (3) along with the homothetic preferences in Eq. (4) allows us to drop k from the set of states and scale variables by k as follows

$$c = C/k, \quad x = X/k, \quad y = Y/k \quad \text{and} \quad g = k'/k, \quad (5)$$

and write the law of motion for k in (2) as

$$g = 1 - \delta + \frac{1}{q}x, \quad (6)$$

and the aggregate resource constraint in (3) as

$$z = c + x. \quad (7)$$

⁹This simplifies to time-separable isoelastic preferences when $\rho = \gamma$.

Law of motion of the shocks.—Let $S = (q, z)$ follow a Markov process with distribution $\psi(S' | S, \theta)$ where θ represents the parameters of the transition law. Some components of θ will be unknown. We will be interested in comparing the behavior of an economy in which agents learn about θ over time to an economy in which θ is known.

Updating of beliefs when there is no regime shift.—A “regime” is denoted by θ . Let $\mu(\theta)$ denote beliefs over θ . Each period the agents observe S_t and update μ using Bayes law. After observing the value S' ,

$$\mu'_{\text{no shift}}(\theta) = \frac{\psi(S' | S, \theta) \mu(\theta)}{\int \psi(S' | S, \theta) d\mu(\theta)} \equiv b(\theta | S', S, \mu), \quad (8)$$

Regime shifts.—We assume that each period a regime shift occurs with probability λ . We expect λ to be close to zero given the above discussion of the probable infrequency of regime shifts. The simplest treatment is to posit a transition law for regimes upon a shift. So we assume that the distribution of the new regime θ' conditional on being in regime θ and conditional on there being a regime shift is $\pi(\theta' | \theta)$ with $\theta \in \Theta$ and π a measure on Θ . In some examples we shall assume that Θ is a finite set so that θ follows a discrete-state Markov process as in Hamilton (1989), Bianchi and Melosi (2016), and Foerster *et al.* (2016).

Timing of a regime shift.—When a shift occurs, it does so at the start of the production period so that agents update by observing S and that $\chi = 1$.¹⁰

Updating of beliefs after a regime shift.—We shall assume that agents know the law $\pi(\cdot)$ and whether a regime shift has just occurred, but that they know neither θ (over which they hold beliefs μ) nor θ' . In that case¹¹

$$\mu'_{\text{shift}}(\theta') = \int \pi(\theta' | \theta) b(\theta | S', S, \mu) d\theta. \quad (9)$$

I.e., S' is generated by the pre-regime-shift θ via the likelihood ψ .

General updating.—Letting $\chi = 1$ denote a regime shift and $\chi = 0$ denote no shift

$$\mu'(\theta) = (1 - I_{\chi=1}) \mu'_{\text{no shift}}(\theta) + I_{\chi=1} \mu'_{\text{shift}}(\theta). \quad (10)$$

¹⁰In Sec. 3.2 we compare to the case where the regime shift dates are not observed.

¹¹The assumption that break dates are known is for tractability. Some of our examples are simple enough to handle unobserved regime shifts (e.g. the finite-state Markov model in section 3.2.2), but others are not. For instance, in the models of section 4, time since the last break is a state variable. If break dates were unobserved, the *distribution* over time since the last break would be a state variable, and the resulting curse of dimensionality would make the calculations intractable. We are still thinking about ways to address this issue.

Plan for the analysis.—Because the model features no monopoly power and has no externalities, the (recursive) competitive equilibrium will coincide with the planner’s solution. Of course the planner too has to learn θ by observing s_t in the same way as private agents do. We shall therefore solve the planner’s problem first, and then discuss the markets for goods, for capital and for assets that decentralize the optimum.

2.1 Planner’s solution

Let μ denote the planner’s belief and S the current shock. Denote the augmented state by $s \equiv (S, \mu)$.¹² The next period’s state is $s' = (S', \mu')$. The distribution of S' is $\int \Psi(S' | S, \theta) d\mu(\theta)$ and it is the same whether a regime occurs or not. The only thing that affected by a regime shift is μ' . Therefore,

$$F(s' | s, \chi') = \begin{cases} F(S', \mu'_{\text{shift}} | s) & \text{if } \chi' = 1 \\ F(S', \mu'_{\text{no shift}} | s) & \text{if } \chi' = 0. \end{cases} \quad (11)$$

so the planner’s predictive distribution is

$$F(s' | s) = \lambda F(S', \mu'_{\text{shift}} | s) + (1 - \lambda) F(S', \mu'_{\text{no shift}} | s), \quad (12)$$

and it underlies the expectations operator in the next two equations and in Appendix B. The planner’s state vector is (s, k) . He chooses C and X to maximize his value function V

$$V(s_t, k_t) = \max_{C_t, X_t} \left[(1 - \beta) C_t^{1-\rho} + \beta \left((E_t[V(s_{t+1}, k_{t+1})^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}}) \right)^{\frac{1}{1-\rho}} \right],$$

subject to (7), (2), (10) and (12).

Because production has constant returns and preferences are homogeneous of degree $1 - \rho$, we can eliminate scale, k , from V and deal in terms of s alone. More precisely, suppose there exists a function $w(s)$ that satisfies the recursion

$$w(s) = \left(\frac{\beta}{1 - \beta} \right)^{\frac{1}{\rho}} \left[E \left((z' + q'(1 - \delta)) \left[(1 - \beta)(1 + q'^{1-\frac{1}{\rho}} w(s'))^\rho \right]^{\frac{1}{1-\rho}} \right)^{1-\gamma} \mid s \right]^{\frac{1-\rho}{\rho(1-\gamma)}}. \quad (13)$$

Proposition 1 • *The value function takes the form*

$$V(s, k) = v(s)k, \quad (14)$$

¹²Since χ is i.i.d. and since beliefs μ update right after χ is observed, including μ as a part of the state means that χ drops out of the state vector.

where

$$v(s) = (z + q(1 - \delta)) \left[(1 - \beta)(1 + q^{1-\frac{1}{\rho}}w(s))^\rho \right]^{\frac{1}{1-\rho}}. \quad (15)$$

- Consumption c takes the form

$$c = \frac{z + q(1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)}. \quad (16)$$

- Growth g takes the form

$$g = 1 - \delta + \frac{zq^{-\frac{1}{\rho}}w(s) - (1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)}. \quad (17)$$

The proof is in Appendix B. The relative simplicity of the value in (15) and the policy rules in (16) and (17) stems from the model's Ak structure. The proposition helps because the right-hand side of (13) contains no max operator; only the level of w and not its derivatives enter the solutions in (15), (16) and (17), and this will help characterize them.

3 A special case

This section specializes the model so as to focus on two specific questions. We assume that

- (i) $\rho = \gamma$ so that utility is the time additive CRRA form $E \sum_0^\infty \beta^t \frac{1}{1-\gamma} c_t^{1-\gamma}$, and
- (ii) S is i.i.d. conditional on θ , i.e., $\psi(S' | S, \theta)$ is independent of S .

We shall analyze parameter learning about the distribution of z separately from that of parameter learning about the distribution of q . I.e., we shall analyze them one at a time. When learning about the distribution of z we shall assume that

$$z = a_z + \frac{A_z}{1 + e^{-x_z}}, \quad (18)$$

where $x_z \sim N(\theta_z, \sigma_z^2)$. The parameters a_z , A_z and σ_z^2 will be known, and θ_z will be unknown. And when learning about the distribution of q we shall assume that

$$q = a_q + \frac{A_q}{1 + e^{-x_q}}, \quad (19)$$

where $x_q \sim N(\theta_q, \sigma_q^2)$, where (a_q, A_q, σ_q^2) are known, and θ_q is unknown. We shall henceforth drop the z and q subscripts from x and from the parameters (a, A, θ) . We now ask two questions:

1. How does a learning-about- θ episode compare to an episode in which θ is known but in which the variance of the shocks x temporarily rises? Sec. 3.1 will show that uncertainty induced by structural breaks has substantially larger and qualitatively different effects than uncertainty triggered by volatility of the exogenous shocks. This is because uncertainty about the parameters of the data generating processes has larger effects on agents' lifetime utility.
2. Second, how does the case where the occurrences of regime shifts, χ , are observed differ from the case when they are not observed? Sec. 3.2 will show that the value of being informed about regime shifts is about twice as large for the case of z as it is for the case of q .

3.1 Parameter learning vs. higher shock volatility

Suppose that a structural break has just occurred, and that agents know this. They also know that thereafter there will be no more breaks, so that $\lambda = 0$. Following the structural break the prior variance over θ rises implying a temporary rise in the variance of the predictive distribution over s .

To put the two cases on a comparable footing, we equate the one-step ahead marginal distributions of the shocks at each t . We wish to compare mean preserving spreads in the distributions of shocks to learning, while keeping the means the same for the two cases. We shall make the comparison for a single learning episode that is not interrupted by a new regime shift ($\lambda = 0$).

The learning economy.—Agents (or the social planner) know the exact values of the parameters $\{\theta_L, \theta_M, \theta_H\}$, but they do not know the regime in place, they have to learn that. Let $h^t = (x_s)_{s=0}^{t-1}$ denote the history of realizations of x ,¹³ and let $L(h^t | \theta_i)$ denote the likelihood of h^t . Assuming that the prior over $\{\theta_L, \theta_M, \theta_H\}$ is $1/3, 1/3, 1/3$, Bayes rule yields the posterior probability that $\theta = \theta_i$ as

$$\mu_i(h^t) = \frac{L(h^t | \theta_i)}{\sum_{j \in \{L, M, H\}} L(h^t | \theta_j)},$$

¹³Agents see z_t or q_t and can recover x_t by inverting (18) or (19) to get

$$x_t = \ln \left(\frac{z_t - a}{A - z_t + a} \right).$$

and the posterior mean and variance as

$$\begin{aligned}\bar{\theta}(h^t) &= \sum \theta_j \mu_j, \quad \text{and} \\ \sigma_{\bar{\theta}}^2(\mu) &= \sum (\theta_j - \bar{\theta}[h^t])^2 \mu_j.\end{aligned}\tag{20}$$

Suppose that the draw following the regime shift is $\theta = \theta_M$, although this will only gradually become apparent to agents as they learn. We condition the distribution of the history on $\theta = \theta_M$, so that $x_t = \theta_M + \varepsilon_t$. The total variance of x is

$$\sigma_{x,t}^2 = \sigma_{\theta,t}^2 + \sigma_{\varepsilon}^2,\tag{21}$$

where $\sigma_{\theta,t}^2$ is the expected variance conditional on θ_M ,

$$\sigma_{\theta,t}^2 = \int \sigma_{\theta}^2(\mu[t, u]) \frac{\sqrt{t}}{\sigma_{\varepsilon} \sqrt{2\pi}} \exp\left(-\frac{(u - \theta_M)^2}{2\sigma_{\varepsilon}^2 t^{-1}}\right) du.\tag{22}$$

The no-learning economy.—In this economy agents *know* that $\theta = \theta_M$, and they face a time-varying variance of ε given by the RHS of Eq. (21), $\sigma_{\varepsilon,t}^2 = \sigma_{x,t}^2$. We shall use the symbol F_t to denote the resulting sequence of time-varying distributions. We treat z learning and q learning one at a time. The treatment of the two is parallel, and we illustrate the treatment of the varying risk value function for the case where z is fixed and only q varies and only q is subject to a possible regime shift. Since the shock variances now depend on t , so does the value:

$$V_t(q, k) = \max_{C, X} \left\{ \frac{C^{1-\gamma}}{1-\gamma} + \beta \int V_{t+1}(q', k') dF_t(q') \right\} = v_t(q) k_t^{1-\gamma}$$

where, using the aggregate resource constraint (7) and law of motion for capital (2),

$$\begin{aligned}v_t(q) &= \max_x \left\{ \frac{(z-x)^{1-\gamma}}{1-\gamma} + \beta \left(1 - \delta + \frac{1}{q}x\right)^{1-\gamma} \int v_{t+1}(q') dF_t(q') \right\}, \\ &= \max_x \left\{ \frac{(z-x)^{1-\gamma}}{1-\gamma} + \beta \frac{\left(1 - \delta + \frac{1}{q}x\right)^{1-\gamma}}{\sigma_{\varepsilon,t} \sqrt{2\pi}} \int v_{t+1}(q') \exp\left(-\frac{(q' - \theta_M)^2}{2\sigma_{\varepsilon,t}^2}\right) dq' \right\}.\end{aligned}\tag{23}$$

For the known- θ case, c_t and g_t are computed directly from equations (13), (16), and (17).¹⁴

¹⁴The Proof Proposition 1 does not impose the constraint that $g \geq 1 - \delta \iff x \geq 0$ (see Eq. (6)) but the inequality always holds in our simulations.

Parameters are calibrated so that the time series means of (c, g) in the learning economy match $c = C/K = 0.3$ and $g = 1.02$ per annum. Results are shown in table 3.1. The same parameters are used for the known case at $\theta = \theta_M$.

Table 3.1: Parameters used for figure 2

| | γ | β | δ | q | z | A | a | θ_L | θ_M | θ_H |
|----------|----------|---------|----------|--------|--------|-------|-------|------------|------------|------------|
| Vary z | 4 | 0.95 | 0.05 | 1 | random | 0.113 | 0.167 | -0.780 | -0.428 | 0.213 |
| Vary q | 4 | 0.95 | 0.05 | random | 0.340 | 0.932 | 1.355 | -0.523 | -0.124 | 0.521 |

The top row of figure 2 compares the two economies when the injected uncertainty is over the distribution of z . After 10 years, cumulative growth is 18.49% in the learning economy and 16.61% in the MPS case. Since the two economies have the same C/K ratios after a decade and K is 1.88 percent higher in the learning economy, it follows that consumption is permanently higher by 1.88 percent.

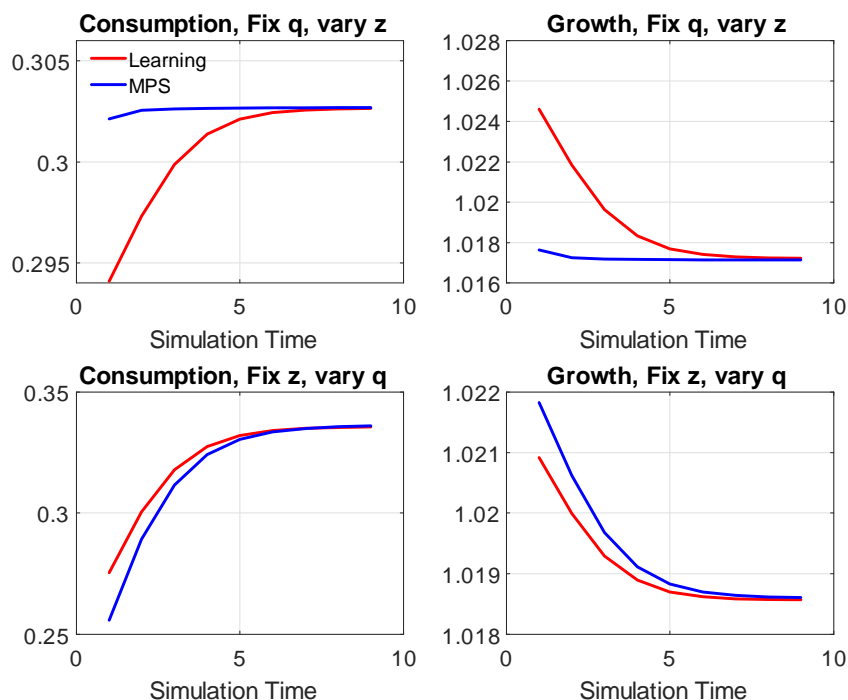


Figure 2: COMPARISON OF LEARNING (RED) AND MPS (BLUE) WHEN THE SHOCK DISTRIBUTION IS UNKNOWN. TOP ROW, z . BOTTOM ROW, q .

The bottom row does the same thing for learning about the distribution of q . In this case, the results go in the opposite direction. Cumulative 10-year growth is

22.22% for the MPS and 21.45% for the learning economy. The difference amounts to a permanent consumption dividend of 0.77 percent per annum for agents in the MPS economy.

Precautionary saving and investment versus the option to wait.—The marginal distributions of the shocks are equated so that $\sigma_{\varepsilon,t}^2 = \sigma_{x,t}^2$, but since beliefs are a Martingale in the learning economy, there is a permanent component to shocks that is absent in the MPS economy. Thus in the top two panels a shock to z has a longer expected duration in the learning case, and to prepare for this agents save more than in the MPS case. In the bottom two panels it is the shock to q that now has a longer expected duration in the learning case, and in Sec. 4 we will find that under learning, precautionary savings dominates the option to wait, but by much less for learning about q than it does for learning about z .¹⁵ In line with that finding, learning now generates more consumption and less growth than does MPS.

3.2 Observed vs. unobserved χ_t when breaks recur

We now assume breaks are recurrent ($\lambda > 0$), and that there are only two regimes $\theta \in \{\theta_L, \theta_H\}$. Agents again know the parameters $\{\theta_L, \theta_H\}$, but they do not know the regime in place.

With just two possible values of θ , beliefs can be summarized by the real number $\mu = \Pr(\theta = \theta_H)$ or, equivalently, by the expectation $E(\theta) = \mu\theta_H + (1 - \mu)\theta_L$. Moreover, we assume $\theta_L = -1$ and $\theta_H = 1$, so that

As in the previous subsection, we shall study regime shifts as shifts in the distribution of one of the shocks alone, keeping the value of the other shock fixed. In light of (18) and (19) for z , the favorable state is θ_H whereas for q , the favorable state is θ_L .

3.2.1 Observed break dates

When a break occurs, each regime is equally likely; in other words, $\pi(\theta' | \theta) = \frac{1}{2}$ for all θ and θ' . Following a regime shift, $E(\theta) = 0$, and then $E(\theta)$ drifts towards the realized $\theta \in \{-1, 1\}$ unless the next regime shift occurs in which case $E(\theta)$ reverts to zero.

Fig. 3 portrays expected consumption and growth for various values of λ . On

¹⁵Compare the top-left panel of figures 7 and 10.

the horizontal axis is $E(\theta)$. At higher λ the belief μ matters less because the regime is more likely to soon change. As λ rises, regime persistence falls and the curves portrayed in the left-hand panels become flatter. A higher λ lowers consumption and raises growth in ‘good’ regimes and does the opposite in ‘bad’ regimes. For z , the ‘good’ θ state is θ_H and consumption and savings (and growth) both rise as beliefs shift towards the regime θ_H . For q the good state is θ_L and so the curves slope down, but the consumption curve again flattens out as λ rises.

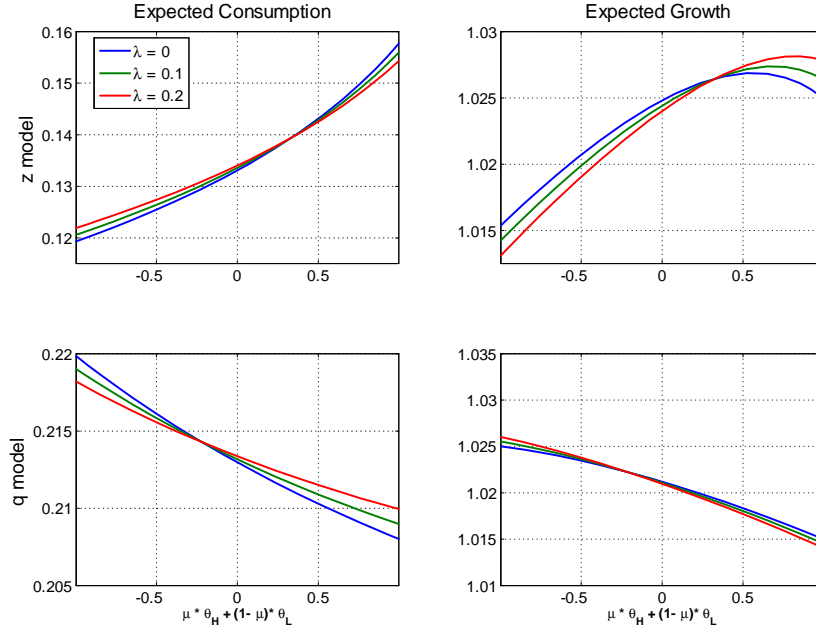


Figure 3: λ , c AND g WHEN χ IS OBSERVABLE. TOP ROW: ONLY θ^z VARIES, BOTTOM ROW: ONLY θ^q DOES.

3.2.2 Unobserved break dates

We now assume the Markov transition probabilities,

$$\theta_z \quad \begin{matrix} & \theta_L & \theta_H \\ \theta_L & \pi_{LL}^z & \pi_{LH}^z \\ \theta_H & \pi_{HL}^z & \pi_{HH}^z \end{matrix}, \quad \theta_q \quad \begin{matrix} & \theta_L & \theta_H \\ \theta_L & \pi_{LL}^q & \pi_{LH}^q \\ \theta_H & \pi_{HL}^q & \pi_{HH}^q \end{matrix},$$

where (θ'_z, θ'_q) are mutually independent.¹⁶ We also allow the structural break indicator χ_t to be observed or unobserved. If the χ_t are not observed, situations may arise in which agents think that a break occurred, while in fact it did not.¹⁷ For this example, we are mainly interested in building intuition about the consequences of assuming that agents know when a break has occurred.

Observability of χ adds nothing if $\lambda \in \{0, 1\}$. When $\lambda = 0$ the regime never changes and when $\lambda = 1$ it changes every period. In these extreme cases, however, there are *de facto* no regime shifts since even if $\lambda = 1$, θ is absorbed into an i.i.d. component of the residual. We therefore derive the consequences for $\lambda \in (0, 1)$.

If χ_t is observable, we have

$$\begin{aligned}\mu'_{\text{shift}}(\theta') &= \sum_{\theta \in \{\theta_L, \theta_H\}} \mu'_{\text{no shift}}(\theta) \pi_{\theta, \theta'} \quad \text{when } \chi = 1, \\ &= \mu_{\text{no shift}}(\theta') \quad \text{when } \chi = 0.\end{aligned}$$

Defining $\mu \equiv \mu(\theta_H)$, it follows that

$$\mu' = (1 - \lambda) b(\theta_H | S', \mu) + \lambda (b(\theta_L | S', \mu) \pi_{LH} + b(\theta_H | S', \mu) \pi_{HH}) \quad (24)$$

Then $w(s, \mu)$ is defined via (8), (9), (12) and (13).

On the other hand, if χ_t is unobserved, beliefs are updated as follows. Agents start the period with a prior μ over $\theta = \theta_H$. An observation S leads to the posterior belief

$$\mu(\theta_H | S, \mu) = \frac{L(S | \theta_H) \mu}{L(S | \theta_H) \mu + L(S | \theta_L) (1 - \mu)} \equiv b(S, \mu).$$

Then next-period belief is

$$\mu' = \lambda (\pi_{HH} b(S, \mu) + \pi_{LH} [1 - b(S, \mu)]) + (1 - \lambda) b(S, \mu) \quad (25)$$

The joint distribution of (μ', s') in Eq. (11) now changes as follows. In $F(s', \mu' | s, \mu)$, χ is no longer a state so that $s = S$. Then $\mu' = b(S, \mu)$ is not random once we condition on (S, μ) because we no longer have μ'_{shift} and $\mu'_{\text{no shift}}$ induced by an observation of χ' , and in Eq. (13) $F(s', \mu' | s, \mu)$ is replaced by

$$F(S' | S, \mu) = \begin{cases} L(S' | \theta_H) & \text{w. prob. } \mu' \\ L(S' | \theta_L) & \text{w. prob. } 1 - \mu' \end{cases}.$$

¹⁶In Appendix C we study the alternating-regime case (in which $\pi_{LL} = \pi_{HH} = 0$) analogous to that of Hopenhayn and Muniagurria (1996).

¹⁷Bianchi and Melosi (2016) study such a scenario, but with a different learning mechanism.

where μ' is given in (25). The decision rules remain as in Eqs. (16) and (17). The unconditional long-run expectation of θ is

$$\hat{E}(\theta) = \frac{\pi_{L,H}}{\pi_{L,H} + \pi_{H,L}}\theta_H + \frac{\pi_{H,L}}{\pi_{L,H} + \pi_{H,L}}\theta_L. \quad (26)$$

Beliefs remain non-degenerate.—Unless θ is forever fixed and that agents know that it is fixed, beliefs will remain non-degenerate. When χ is unobserved the agents never learn which of the two values of θ is operative: Beliefs remain a non-degenerate stochastic process even if (unbenownst to the agents) θ remained forever fixed at θ_H , say. If agents' beliefs placed all weight on θ_H , say, at $t - 1$, they would become non-degenerate at t , because of the positive probability $\lambda\pi_{H,L}$ that θ will have shifted due to a possible regime shift. In other words, although they know the set $\{\theta_L, \theta_H\}$, agents will typically not know which regime θ is in place.

For example, suppose parameters are fixed as in table 3.2.2. These parameters are based on the following accounting: The NBER data indicate 215 expansions, of which 204 (94.88%) are followed by an expansion, and 11 (5.1%) are followed by a recession. On the other hand, there are 41 recessions, of which 11 (26.8%) are followed by an expansion, and 30 (73.1%) are followed by another recession. We use the notation $\pi_{L,H}$ and $\pi_{H,L}$ for both cases – learning about z and learning about q since here they are obtained by the same accounting method.

Table 3.2.2: Parameters chosen for figure 4

| θ_L | θ_H | $\pi_{L,H}$ | $\pi_{H,L}$ | λ |
|------------|------------|-------------|-------------|-----------|
| -1 | 1 | 0.27 | 0.05 | 0.29 |

With these parameters, $\hat{E}(\theta) = 0.68$, which is indicated by the vertical lines in figure 4. Not observing χ means that the planner acts on the basis of less information than in the observable- χ case. This appears to imply more uncertainty over the future realizations of (z, q) , and we expected to see a rise in precautionary savings in response. For most values of μ , this occurs when learning about shifts in q (see the bottom row) but not for shifts in z (shown in the top row). The learning-about- z model has the surprising outcome that for most values of μ , agents consume more when regime switches are not observed.

3.2.3 The value of being able to observe χ

As noted above, the ability to observe χ_t is of no value if $\lambda \in \{0, 1\}$, because then one would know that the regime never changes (when $\lambda = 0$) or that it changes every

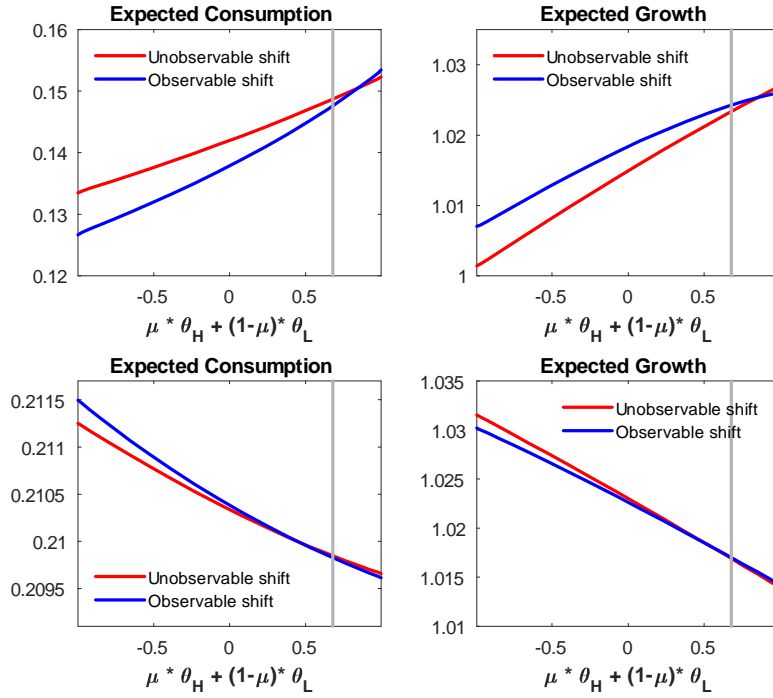


Figure 4: λ , c AND g WHEN χ IS UNOBSERVABLE. TOP ROW: ONLY θ^z VARIES, BOTTOM ROW: ONLY θ^q DOES.

period (when $\lambda = 1$). Thus we expected the value of information to be highest for values of λ around $1/2$.

Figure 5 shows, however, that the value of information is seemingly maximized at values of λ well below $1/2$, and closer to a value of $1/3$ or even $1/4$, depending on the value γ . This is probably because, with $\hat{E}(\theta) = 0.68$, beliefs about θ drift upwards and the value of knowing about regime changes becomes less valuable. The rate of the decline in the value of information over time is increasing in λ .

Figure 5 portrays the relative value of information $v^I(s, \mu) / v^U(S, \mu)$ as λ varies. The experiment fixes $\mu = 0.5$ and sets $z = 0.34$ and $q = 1$. With $\gamma > 0$, $v^I(s, \mu)$ and $v^U(s, \mu)$ are both negative. Then $v^I(s, \mu) > v^U(s, \mu) \implies v^I(s, \mu) / v^U(S, \mu) < 1$. At $\lambda = 0$ and 1 , the ratio is equal to 1 . For intermediate values of λ , being informed about regime shifts raises $v(s, \mu)$. Provided that $\lambda \in (0, 1)$, the value of information rises with the degree of risk aversion.

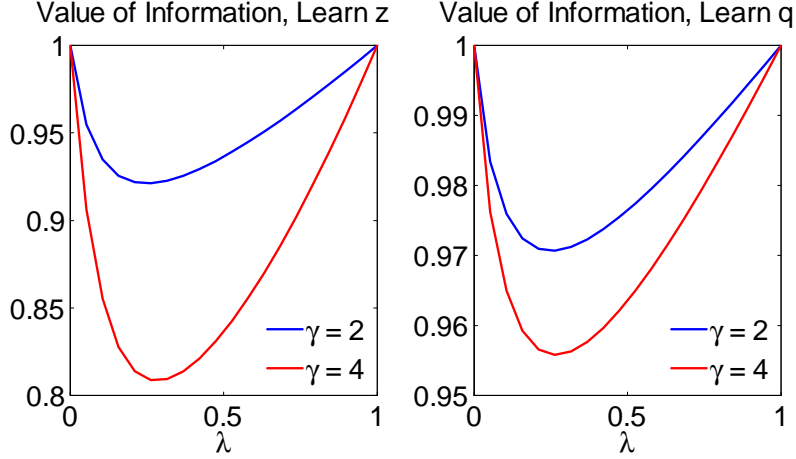


Figure 5: VALUE OF INFORMATION ABOUT REGIME SHIFTS

4 A model with non-recurrent structural breaks and autoregressive shocks

Next we turn to examples in which shocks are autoregressive processes,

$$\begin{aligned}\ln z_t &= \mu_{zt} + \rho_z \ln z_{t-1} + \sigma_z \varepsilon_{zt}, \\ \ln q_t &= \mu_{qt} + \rho_q \ln q_{t-1} + \sigma_q \varepsilon_{qt}.\end{aligned}\tag{27}$$

To keep things simple, we activate one at a time. The first example features neutral technology shocks with a constant level of q , while the second fixes z and activates investment shocks. Letting a_t represent the active shock, the autoregressive parameter ρ_a and conditional standard deviation σ_a are assumed to be known, and the intercept μ_{at} is unknown. Agents must learn about μ_{at} .

The intercept is subject to occasional structural breaks,

$$\begin{aligned}\mu_{at} &= \mu_{at-1} \quad w/ \text{pr } 1 - \lambda, \\ &= m_t \sim N(m, \sigma_m^2) \quad w/ \text{pr } \lambda.\end{aligned}\tag{28}$$

A Bernoulli random variable χ_t governs whether a structural break occurs. With probability $1 - \lambda$, no break occurs ($\chi_t = 0$), and the intercept remains unchanged. With probability λ , a break occurs ($\chi_t = 1$), and a new intercept is drawn from a normal distribution. The random variables $\varepsilon_{at}, \chi_t, m_t$ are mutually independent, and the parameters $\theta = [\lambda, m, \sigma_m, \rho_a, \sigma_a]$ are known.

Agents observe χ_t and a_t , but not μ_{at} or m_t . Because χ_t is observable, agents know when a break occurs, but they don't know whether μ_{at} has increased or decreased or by how much. They update beliefs about μ_{at} by applying Bayes theorem.

To fit this example into our general framework, let $s = [a, \chi]$ represent the observable states, and write the Markov transition equation as $\Psi(s_{t+1}, \mu_{t+1} | s_t, \mu_t, \theta)$. The predictive density is

$$F(s_{t+1} | s^t, \theta) = \iint \Psi(s_{t+1}, \mu_{t+1} | s_t, \mu_t, \theta) p(\mu_{t+1}, \mu_t | s^t, \theta) d\mu_{t+1} d\mu_t, \quad (29)$$

where $p(\mu_{t+1}, \mu_t | s^t, \theta)$ is the posterior for the unobserved intercept. Proposition 1 goes through with this specialization of notation. All that remains is to derive the posterior $p(\mu_{t+1}, \mu_t | s^t, \theta)$ and solve the integral equation 13. Appendix D addresses these problems.

4.1 Neutral technology shocks

Table 4.1 calibrates parameters for an economy in which the neutral technology shock z is active and q is constant. The discount factor β is set *a priori* to a standard RBC value for annual data, and the coefficient of relative risk aversion γ is set to 4. Consumers are myopic if $\gamma = 1$, and a higher value is needed for learning to matter. However, Pratt-style thought experiments suggest that γ should not be much higher than 1. A value of 4 seems like a reasonable compromise.

Table 4.1: Calibration for structural breaks in *TFP*

| β | γ | ρ | δ | ρ_z | σ_z | λ | m | σ_m | \bar{z} | \bar{q} |
|-------------|----------|--------|----------|----------|------------|-----------|---------|------------|-----------|-----------|
| 1.01^{-4} | 4 | 12.95 | 0.119 | 0.8867 | 0.0422 | 0.0091 | -0.0893 | 0.0234 | 0.464 | 1 |

Parameters governing $\ln z$ are estimated by combining the data shown in figure 1 with informative priors over $(\rho_z, \sigma_z, \lambda, m, \sigma_m)$. In a nutshell, the process for $\ln z$ can be expressed as a non-Gaussian state-space model whose log likelihood function can be evaluated via a particle filter. Because structural breaks are infrequent, λ and σ_m are weakly identified in a frequentist sense. We therefore add an informative prior and maximize the log posterior. Details can be found in appendix E. Table 4.1 reports the posterior mode for $(\rho_z, \sigma_z, \lambda, m, \sigma_m)$, and \bar{z} is the implied unconditional geometric mean. Since no model for $\ln q$ is on the table at this point, \bar{q} is set to 1.

Even with structural breaks, $\ln z$ is a stationary random process, although it would be hard to distinguish from an integrated process in samples of 100 or 200 years. That breaks are rare means that agents face long-run risks analogous to those of Bansal and Yaron (2004). The left hand column of figure 6 illustrates the long-run risk by comparing unconditional distributions for processes with and without structural breaks.¹⁸ When structural breaks are absent (solid lines), $\ln z$ is unconditionally normal with mean $m/(1 - \rho_z)$ and variance $\sigma_z^2/(1 - \rho_z^2)$. With rare structural breaks ($\lambda = 0.0091$, dashed lines), the distribution is still centered on $m/(1 - \rho_z)$, but the tails are fatter. In fact, in the tails, the log density for the rare-break process is 10 times greater than that of the no-break process.

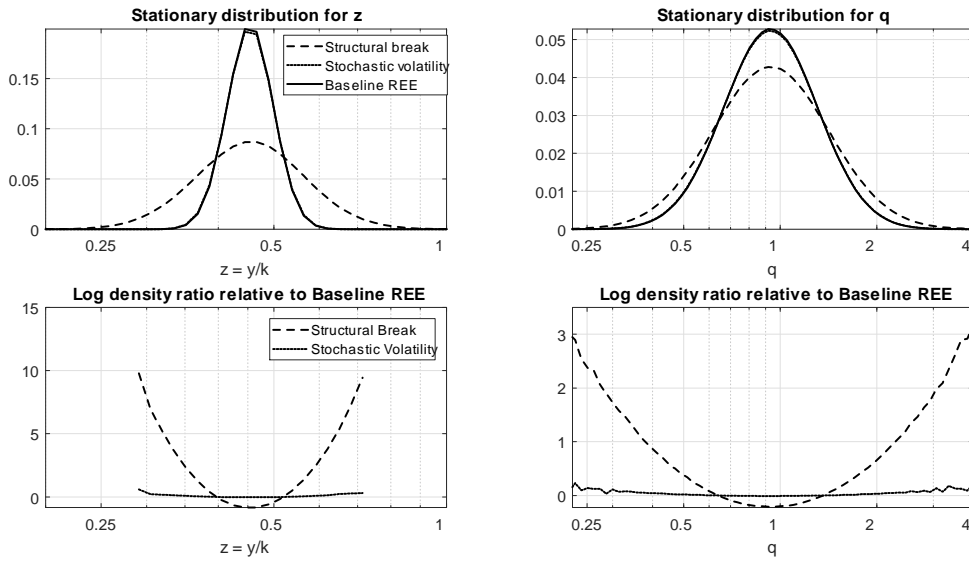


Figure 6: Unconditional shock distributions

The remaining parameters ρ and δ are calibrated by matching aspects of the deterministic steady state.¹⁹ When all shocks are deactivated ($\sigma_z = \lambda = 0$), steady-state growth and the consumption-income ratio are

$$g_d = \beta^{1/\rho} \left(\frac{\bar{z}}{\bar{q}} + 1 - \delta \right)^{1/\rho}, \quad (30)$$

$$\left(\frac{C}{Y} \right)_d = 1 + (1 - \delta - g_d) \frac{\bar{q}}{\bar{z}}.$$

¹⁸The $\ln q$ calibration and stochastic volatility model are discussed below.

¹⁹Results are similar if ρ and δ are calibrated to the mean of a stochastic equilibrium in which shocks to $\ln z$ are active but structural breaks are not ($\sigma_z \neq 0, \lambda = 0$).

The parameters ρ and δ are chosen to match $g_d = 1.02$ and $(C/Y)_d = 0.7$.²⁰ The solution for $\delta = 0.119$ is close to the standard RBC calibration for annual data. The solution for $\rho = 12.95$ implies that the elasticity of intertemporal substitution (EIS) is 0.077 which is broadly consistent with estimates based on micro data (e.g., Hall 1988, Vissing-Jorgensen 2002, Yogo 2004, and Havranak 2015). For instance, in a meta-analysis of 169 publications, Havranak (2015, p.1196) reports a ‘best practice’ estimate of $1/\rho = 0.33$ with a confidence interval of (-0.2,0.8). It follows that our consumers have a strong preference for smooth consumption streams.

The EIS is lower than typical calibrations in the finance literature; e.g., Bansal and Yaron (2004) assume a value of $1/\rho = 1.5$. As they emphasize, a model with EZW preferences requires both long-run risk and an EIS greater than 1 to resolve asset pricing puzzles. Structural breaks generate long run risk, but our model could not hit the macro targets $g_d = 1.02$ and $(C/Y)_d = 0.7$ if the EIS were greater than 1. Furthermore, attempts to calibrate to the risk-free rate plus one of the macro targets were unsuccessful, with moment conditions either failing to solve altogether or resulting in implausible values for other parameters. We chose the calibration in table 4.1 so that the model generates plausible macro outcomes, but we acknowledge that the model will not match asset prices with an EIS close to zero.

As a point of departure, we deactivate structural breaks and parameter uncertainty and compute the *REE* for a conventional *AK* model. The black lines in the top row of figure 7 depicts policy functions for C/Y and g when μ_z is constant and known with certainty. As expected, the consumption-output ratio is decreasing in $\ln z$, and growth is increasing. The former reflects consumption smoothing, because a low current value of $\ln z$ signals higher future productivity. The latter just reflects that an increase in the investment share raises growth in an *AK* model.

The other curves in the top row portray slices of the policy functions when structural breaks and learning are active. This model has three state variables, $\ln z$, the posterior mean $\mu_{t|t}$, and the posterior precision $P_{t|t}$. In the top row, the posterior mean $\mu_{t|t}$ is held constant at the true μ_z , and precision $P_{t|t}$ is indexed by time since the last break, with uncertainty being highest in year 1 and diminishing with t .

²⁰The target for the consumption-income ratio is based on US data for the period 1950-2015. Consumption is measured by real expenditures on nondurable goods and services, and investment is real gross private domestic investment plus expenditures on consumer durables. Because the model has no government or net exports, output is measured by $C + I$. If government spending were included and split between consumption and investment, the mean C/Y ratio would be around 0.75.

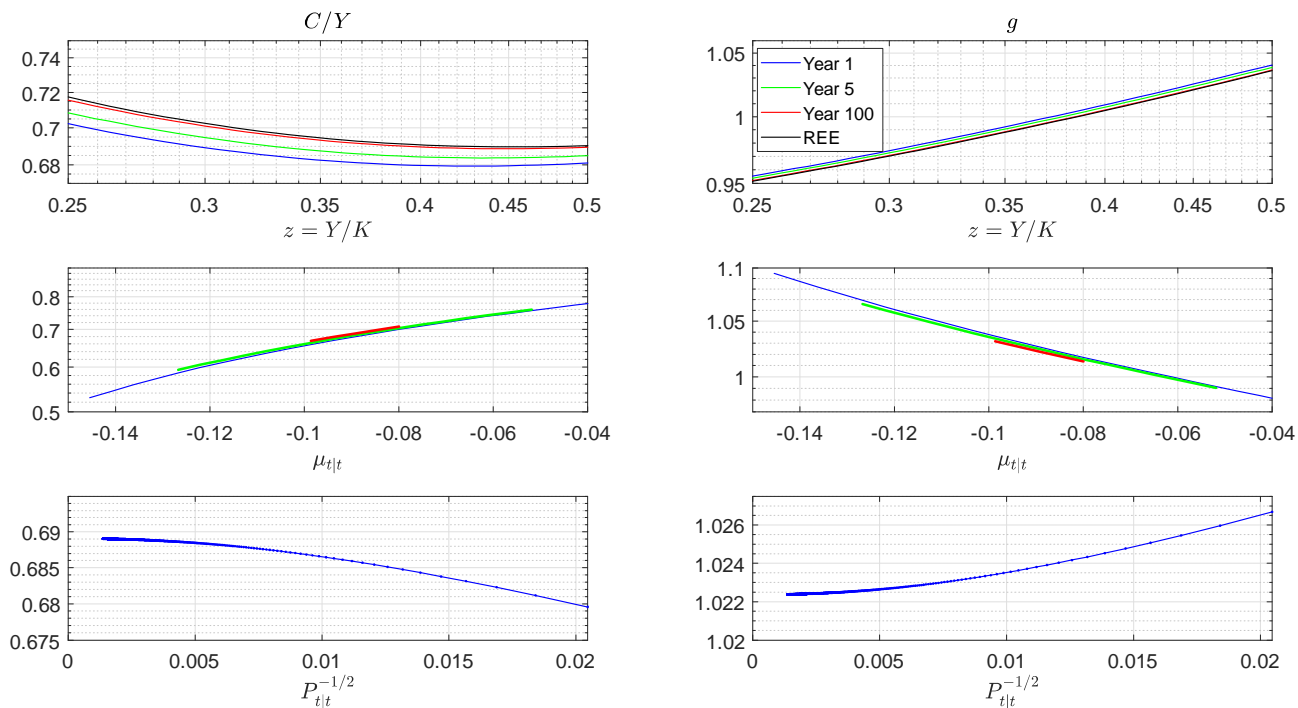


Figure 7: POLICY FUNCTIONS, Z ACTIVE AND Q INACTIVE.

Uncertainty promotes precautionary saving and shifts the policy function for C/Y downward. Because capital is the only aggregate store of wealth, the additional savings are channeled into investment, and g increases. As parameter uncertainty diminishes, precautionary savings decline, and the policy functions shift toward those for the benchmark no-break model.

Since the model has complete markets, a safe asset is implicit. One might wonder why agents don't park their precautionary savings there to await more information. The answer is that the safe asset is in zero net supply and hence is not a store of aggregate wealth. The risk-free rate and return on capital adjust so that agents do not park precautionary savings in the safe asset. In principle, inventories or other durable commodities could serve as an aggregate store of precautionary wealth, but they are not riskless.

The second row illustrates how variation in point estimates $\mu_{t|t}$ influence C/Y and g with $\ln z$ held constant at its unconditional mean. We limit attention to plausible ranges by plotting $\mu_{t|t}$ over intervals of ± 3 posterior standard deviations

around the true value. Variation in $\mu_{t|t}$ matters a lot, especially shortly after a break when point estimates can bounce all over the place. Conditional on no future break, $\mu_{t|t}$ is a martingale under the subjective probability law (Doob 1948), so a revision of the point estimate shifts the agent’s long run forecast for productivity. Hence, consumption responds more strongly than to a transitory innovation ε_{zt} . Although the estimates eventually settle down as t increases, this channel continues to amplify volatility for a long time.

The magnitude of precautionary effects can be seen more clearly in the third row. Here, C/Y and g are depicted as functions of the posterior standard deviation $P_{t|t}^{-1/2}$ with the posterior mean $\mu_{t|t}$ and technology shock $\ln z$ held constant at their unconditional means. The peak effect occurs in the year of a break (shown at the far right of each panel). The consumption-income ratio is about 1 percentage point lower than in the benchmark no-break model, and growth is about 40 basis points higher. As uncertainty is resolved (moving from right to left), the investment share falls and growth slows. It takes about 5 years to move halfway across the graph and 20 years to move two-thirds of the way. Beyond that point, the policy functions become flatter, and the effects of parameter uncertainty diminish. Most of the precautionary effects occur in the first two decades. The effects of variation in $\mu_{t|t}$ persist beyond that point, however.

Further insight can be obtained by simulating the model. Figure 8 compares outcomes for economics with identical sequences of (scaled) *TFP* shocks. The simulations are initialized by setting $\ln z_0$ at its unconditional mean, drawing 10,000 sample paths for $\ln z_t$, and then calculating outcomes for c, g by plugging the simulated shocks into the policy functions. Blue lines in the top row portray the median and interquartile range at each date for the structural-break model, while black lines depict those for the baseline *REE*. The second row depicts differences across sample paths on which standardized realizations of ε_{zt} are held constant.

In addition to conventional innovations in *TFP* (ε_{zt}), the learning economy is also subjected to a ‘pure uncertainty’ shock at date zero. To create a pure uncertainty shock, we posit a structural break in which the newly drawn value of μ_z happens to coincide with the old value and the prior mean; i.e., $\chi_t = 1$, $m_t = \mu_{t-1} = m$. In other words, a pure uncertainty shock is fake news: agents believe a structural break has occurred when in fact it has not.

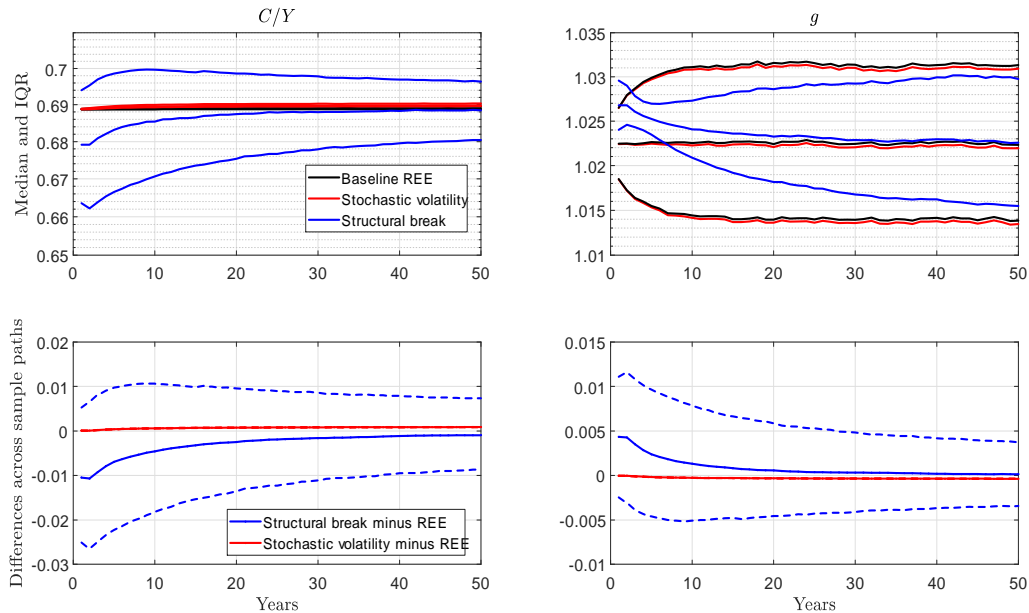


Figure 8: FAN CHARTS, Z ACTIVE AND Q INACTIVE. IN THE FIRST ROW, BLUE LINES PORTRAY THE MEDIAN AND INTERQUATILE RANGE FOR A PURE UNCERTAINTY SHOCK, WHILE BLACK AND RED LINES DEPICT THOSE FOR THE BASELINE REE AND A REE WITH STOCHASTIC VOLATILITY, RESPECTIVELY. THE SECOND ROW DEPICTS DIFFERENCES BETWEEN THE LEARNING AND STOCHASTIC VOLATILITY MODELS RELATIVE TO THE BENCHMARK REE MODEL, WITH SOLID LINES SHOWING THE MEAN AND DASHED LINES THE INTERQUARTILE RANGE.

A structural break activates two belief states, $\mu_{t|t}$, and $P_{t|t}$. A fake news shock creates uncertainty about μ_z , raising the posterior variance $P_{t|t}^{-1}$ and increasing the sensitivity of the conditional mean $\mu_{t|t}$ to incoming data. Their combined effect is to depress the average consumption share, raise average growth, and amplify consumption volatility.

The precautionary effects can be seen most clearly in the second row. Solid lines portray cross-sectional average differences between the structural-break model and the no-break economy. Because no actual break has occurred, averaging across sample paths removes the effects of $\mu_{t|t}$ and isolates the influence of $P_{t|t}$.²¹ Furthermore, because the posterior variance is a deterministic function of time since the last break, it is the same on all sample paths. The cross-sectional average therefore illustrates

²¹I.e., $\mu_{t|t}$ averages to μ_{zt} across sample paths. Since there are no actual breaks, μ_{zt} remains constant at m .

how higher uncertainty about μ_z influences C/Y and g with the mean of $\mu_{t|t}$ held constant at the true μ_z . The peak differences occur in the first 5 years, when the mean consumption share is 0.8 to 1.1 percentage points lower and mean growth is 25 to 45 basis points higher. By the end of the second decade, the effects of μ_z uncertainty are weaker, but the risk of a future break remains. Consequently, the mean consumption share remains about 25 basis points lower than in the baseline *REE*, and mean growth is about 5 basis points higher. These small growth effects continue to add up, and after 50 years the mean capital stock is 4.75 percent higher than in the *REE*.²²

The structural-break model also features greater consumption volatility, due largely to variation in the conditional mean $\mu_{t|t}$. Because $\mu_{t|t}$ evolves as a martingale (conditional on no future break), it has a strong influence on permanent income. Hence variability in $\mu_{t|t}$ strongly amplifies consumption volatility. For a model like ours, Jones, et al. (2005) report that consumption would be too smooth under *REE* for an *EIS* as low as ours. This is also true in our model (the black fan chart in the upper left panel is so narrow that it looks like a single line). When $\mu_{t|t}$ is active, however, the model exhibits plenty of consumption volatility despite the low *EIS* (see the blue fan chart in the upper left panel). Indeed, on many sample paths, the pure precautionary effect is swamped by movements in permanent income. In other words, while average C/Y is lower in the learning model, there are many sample paths on which it is higher than in the *REE* (compare the means and interquartile ranges in the second row). Of course, variation in $\mu_{t|t}$ is itself a consequence of the higher uncertainty because $\mu_{t|t}$ would be less sensitive to incoming data if $P_{t|t}^{-1}$ did not rise.

Since a pure uncertainty shock is somewhat contrived in this environment, we also look at the combined effects a structural break along with the associated increase in uncertainty. Figure 9 illustrates the differences between a fake news shock and breaks of magnitude $\pm 1\sigma_m$. As in the second row of figure 8, the solid lines represent mean differences across sample paths, with realizations of ε_{zt} held constant, while the dashed lines represent the interquartile range.

After a signal that $\chi_{t+1} = 1$, the posterior variance rises by the same amount for both shocks, and the precautionary effects net roughly to zero when differencing

²²A higher *EIS* would amplify the response of precautionary savings because consumers would be more willing to trade lower present consumption for higher future consumption. For instance, for $1/\rho = 0.33$ (Havranak's point estimate), the mean consumption share would be 140 basis points lower than in the *REE* at impact, mean growth would be 55 basis points higher, and the capital stock would be 8 percent higher after 50 years.

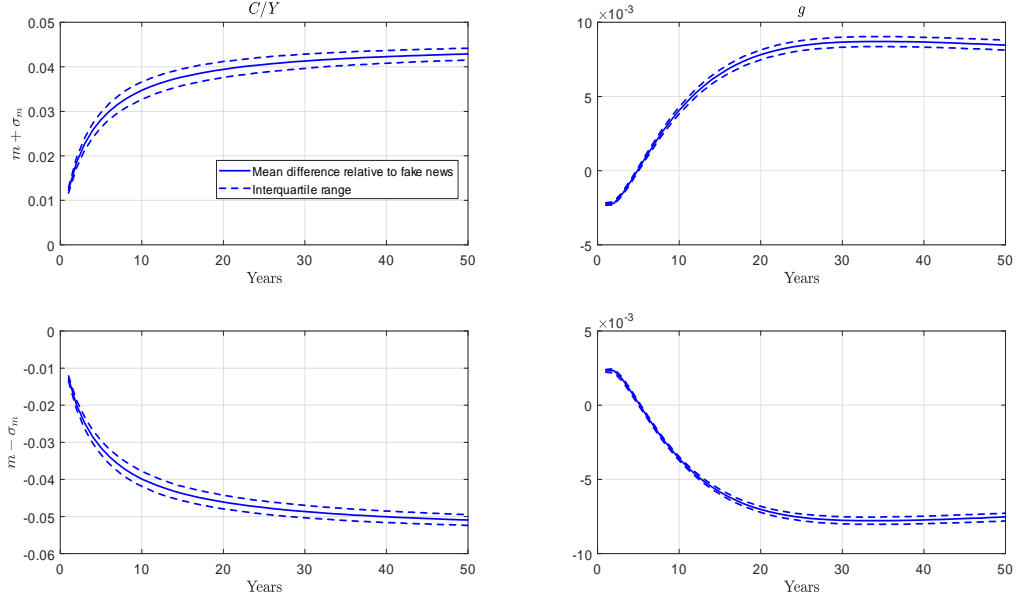


Figure 9: Fake news about μ_{z_t} versus actual breaks

across sample paths. What remains are the effects of $\mu_{t|t}$. When μ_z rises (top row), agents attribute the increase in productivity partly to an increase in μ_z and partly to a positive innovation in ε_z . The former raises the consumption share, while the latter depresses it. Relative to a fake news shock, more of the increase in $\ln z_t$ is attributed to μ_z . As a consequence, the net impact effect on the mean C/Y ratio is positive. As agents gradually verify that μ_z is actually higher, both C/Y and g rise. After 50 years, the mean consumption share is 4 percentage points higher than for a fake news shock, average growth is about 90 basis points higher, and the mean capital stock is about 30 percent higher. The effects of a downward shift in μ_z are similar but have the opposite sign (bottom row).

Last but not least, we contrast uncertainty about μ_z with uncertainty about future innovations $\varepsilon_{z,t}$. We do this by introducing stochastic volatility in a model in which μ_z is constant and known with certainty. In the structural-break model, the prediction error variance for $\ln z_t$ is

$$\begin{aligned} \text{var}(\ln z_{t+1} | \ln z^t, \chi_{t+1} = 1) &= \sigma_z^2 + \sigma_m^2, \\ \text{var}(\ln z_{t+1} | \ln z^t, \chi_{t+1} = 0) &= \sigma_z^2 + P_{t|t}^{-1}, \end{aligned} \quad (31)$$

where t indexes time since the last volatility break. To put a stochastic volatility model on the same footing, we assume that the conditional variance for ε_{zt} jumps to $\sigma_z^2 + \sigma_m^2$ with probability λ and then declines with probability $1 - \lambda$ to $\sigma_z^2 + P_{t|t}^{-1}$. The red lines in figure 8 portray outcomes for the stochastic volatility model.

Adding stochastic volatility to the *REE* model matters slightly, but does not change the big picture. This is especially clear in the second row, where red lines depicts differences across sample paths relative to the *REE* economy. When shown on the same scale as sample-path differences for the structural-break economy, those for the stochastic-volatility model are too small to see. The reason why structural breaks are more important is that they generate more persistent variation in the conditional mean. As a consequence, they have a greater impact on permanent income.

Furthermore, stochastic volatility creates less long-run risk. As shown in figure 6, structural breaks fatten the tails of the unconditional distribution of $\ln z$ more. Indeed, the long-run risk depicted there depends mainly on variation in the conditional mean. Matching the conditional variance of ε_{zt} to the prediction-error variance of the structural-break process fattens the tails on the unconditional distribution only slightly.

To summarize, structural breaks have big effects in this model, as do movements in the conditional mean $\mu_{t|t}$. Movements in the conditional variance $P_{t|t}^{-1}$ matter, but their effects are smaller in magnitude. Smaller still are the effects of stochastic volatility in ε_{zt} .

4.2 Investment-specific technology shocks

Results for a model in which q is active and z is inactive are qualitatively similar. An uncertainty shock promotes savings and growth, and the responses are larger in magnitude than those associated with a comparable increase in conditional variances.

Parameters for this model are shown in table 4.2. The subjective discount factor β , coefficient of relative risk aversion γ , and steady-state value of z are the same as in the previous section. The parameters governing $\ln q_t$ are estimated by maximizing the log posterior for the data shown in figure 1, with priors as described in appendix E.3. Given these parameters, the inverse-EIS ρ and depreciation rate δ are chosen so that the deterministic steady states for C/Y and g are 0.7 and 2 percent per annum, respectively. Once again, the implied EIS is close to zero.

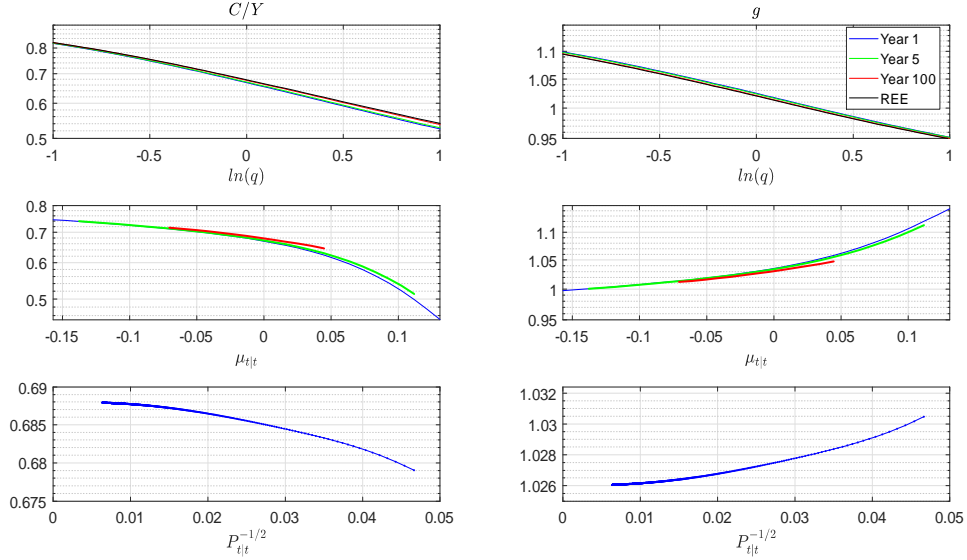


Figure 10: POLICY FUNCTIONS, Q ACTIVE AND Z INACTIVE.

Table 4.2: Calibration

| β | γ | ρ | δ | ρ_q | σ_q | λ | m | σ_m | \bar{z} | \bar{q} |
|-------------|----------|---------|----------|----------|------------|-----------|--------|------------|-----------|-----------|
| 1.01^{-4} | 4 | 13.8045 | 0.129 | 0.81 | 0.198 | 0.0111 | -0.013 | 0.048 | 0.464 | 0.934 |

The black lines in the top row of figure 10 portray policy functions for a version that abstracts from breaks and learning. As q increases, the capital-goods technology becomes less efficient, and each unit of investment is transformed into fewer units of new capital. To compensate, agents feed more investment goods into the capital-goods sector. Thus, the consumption share decreases with q . The rise in the investment share only partly offsets the decline in the efficiency of the capital-goods technology, however, implying that the capital stock also grows more slowly as q increases.

The colored lines in the top row illustrate the effects of uncertainty. In this panel, the posterior mean $\mu_{t|t}$ is held constant at the true μ_q , and precision $P_{t|t}$ is indexed by time since the last break, with uncertainty being highest immediately after a structural break. The qualitative effects are the same as for a neutral *TFP* shock: uncertainty promotes precautionary saving, increases the investment share, and shifts the policy function for g upward. As uncertainty is resolved, precautionary savings decline, and the policy functions shift toward those for the no-break benchmark.

The middle row illustrates how variation in $\mu_{t|t}$ influences C/Y and g with $\ln q$ held constant at its unconditional mean. Attention is limited to plausible values by plotting $\mu_{t|t}$ over intervals of ± 2 posterior standard deviations around the true value. Variation in $\mu_{t|t}$ matters a lot, especially shortly after a break when point estimates are most sensitive to realizations of $\ln q_t$. Recall once more that $\mu_{t|t}$ is a martingale conditional on no break, so the agents's (conditional) long run forecast for $\ln q_t$ reacts strongly to this state variable. Hence, consumption and growth also respond strongly. This effect is nonlinear, being largest for high values of $\mu_{t|t}$.

The third row isolates the influence of $P_{t|t}$. Here C/Y and g are plotted as functions of the posterior standard deviation $P_{t|t}^{-1/2}$ with $\ln q$ and $\mu_{t|t}$ held constant at their respective unconditional means. Because $P_{t|t}^{-1/2}$ declines deterministically with time since the last break, time runs from right to left in this row. The points closest to the origin illustrate the effects of background uncertainty associated with conventional shocks to $\ln q$ and potential future breaks. Since $C/Y = 0.7$ and $g = 1.02$ in the deterministic steady state, background uncertainty reduces the consumption share and raises growth by roughly 120 and 60 basis points, respectively. The points on the far right depict the impact effects of an uncertainty shock. In the year of a structural break, the consumption share falls by another 90 basis points and growth increases by a further 45 basis points. As uncertainty is resolved, the investment share declines back toward its stochastic steady state, and growth declines.

The next figure compares outcomes for economies with and without structural breaks. There are two of the latter, one in which the conditional variance of ε_{qt} is constant and another in which it is calibrated to match the prediction-error variance of the structural-break model (cf. the discussion surrounding equation 31). To create a pure uncertainty shock, we again imagine that agents believe a structural break has occurred when in fact it has not. As before, the three simulations used the same scaled ε_{qt} shocks, so shocks are held constant on all sample paths.

In the top row, blue lines portray the median and interquartile range for the structural-break model, while black and red lines depict those for the benchmark *REE* and stochastic-volatility models. The fact that the black and red lines lie on top of one another implies that stochastic volatility matters very little when μ_q constant and known with certainty. In contrast, in response to a pure uncertainty shock, the median consumption share falls by about 100 basis points and then rises gradually as uncertainty is resolved. The median growth rate rises by about 45 basis points at

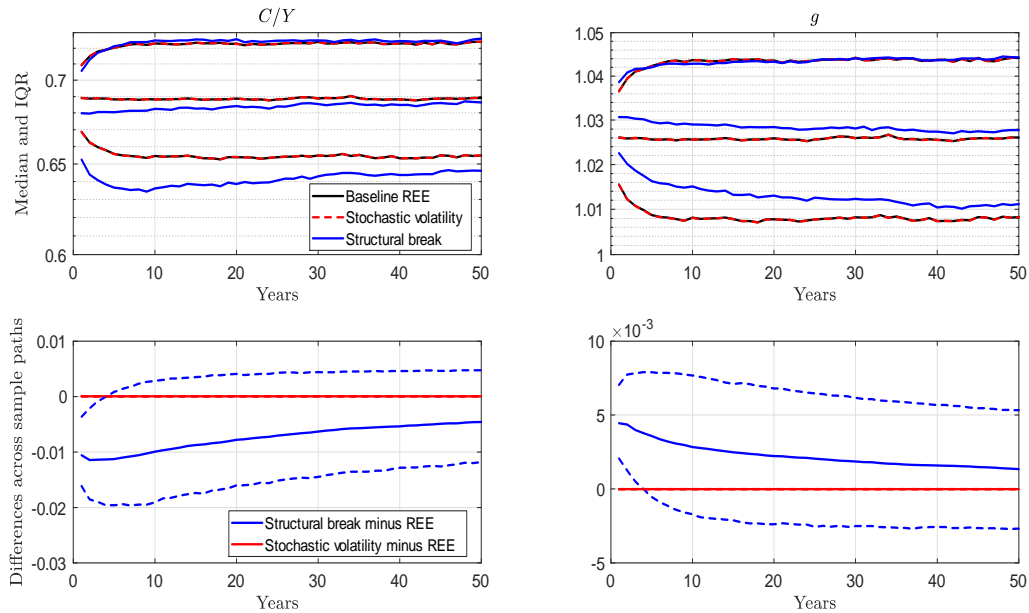


Figure 11: FAN CHARTS, Q ACTIVE AND Z INACTIVE. IN THE FIRST ROW, BLUE LINES PORTRAY THE MEDIAN AND INTERQUATILE RANGE FOR A PURE UNCERTAINTY SHOCK, WHILE BLACK AND RED LINES DEPICT THOSE FOR THE BASELINE REE AND A REE WITH STOCHASTIC VOLATILITY, RESPECTIVELY. THE SECOND ROW DEPICTS DIFFERENCES BETWEEN THE LEARNING STOCHASTIC VOLATILITY MODELS RELATIVE TO THE BENCHMARK REE MODEL, WITH SOLID LINES SHOWING THE MEAN AND DASHED LINES THE INTERQUATILE RANGE.

impact, and declines slowly.

The second row highlights differences between the economies. Solid lines portray cross-sectional average differences relative to the *REE* and stochastic-volatility economies, while dashed lines depict the interquartile range. As before, the cross sectional average illustrates how higher uncertainty influences C/Y and g with the mean of $\mu_{t|t}$ held constant at the true μ_q . For the structural-break model, the biggest differences occur in the first 5 years, when the mean consumption share is about 100 basis points percentage points lower and mean growth is 30 to 50 basis points higher. By the end of the second decade, the effects uncertainty are weaker, but the risk of a future break remains. Consequently, the mean consumption share remains about 80 basis points lower than in the *REE*, and mean growth is about 25 basis points higher. Although this might not seem like much, the cumulative effect over 50 years is an 11 percent increase in the level of the capital stock.

As before, when breaks are absent, outcomes for a stochastic-volatility economy are essentially the same as for the benchmark *REE* model with constant conditional variances. Differences across sample paths are too small to see when graphed on the same scale as the structural-break economy. Structural breaks matter more because they activate persistent variation in the conditional mean of $\ln q$.

Finally, figure 12 illustrates the differences between structural breaks of magnitude $\pm 1\sigma_m$ and a fake news shock. Solid lines again represent mean differences across sample paths, with realizations of ε_{qt} held constant, while the dashed lines represent the interquartile range.

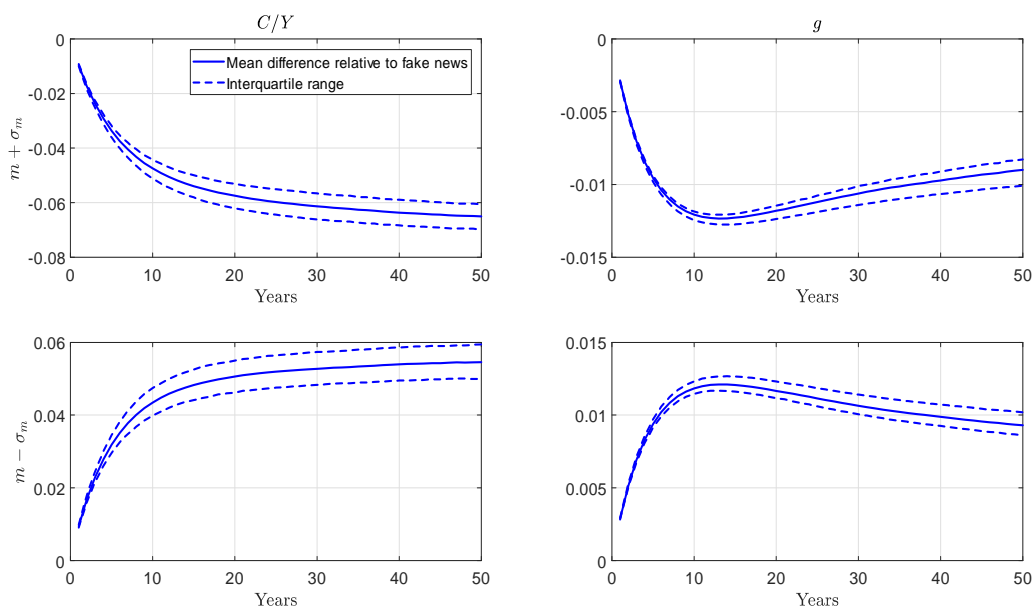


Figure 12: Structural breaks in μ_q versus. fake news

Since the effects of positive and negative structural breaks are approximately symmetric, it suffices to discuss the effects of a structural break that increases μ_q (top row). Recall that the posterior variance rises by the same amount for actual breaks and fake news, implying that the precautionary effects net roughly to zero when differencing across sample paths. Thus, the differences primarily reflect the influence of $\mu_{t|t}$. After both shocks, agents attribute the decrease in capital-goods productivity partly to an increase in μ_{qt} and partly to a positive innovation in ε_{qt} , but more to μ_{qt} when there is an actual break. After an actual break, as agents gradually verify that μ_{qt} is higher, the decline in C/Y is magnified. After 50 years,

the mean consumption share for an actual break is 6 percentage points lower than for a fake-news shock. The effects on g are similar, but with some over-shooting in the medium run.

5 Conclusion

This paper analyzed learning by agents of two aggregate-shock distributions. The economy is competitive and has no external effects. We solved the planner's problem and then explained how decentralization would effect the same outcome.

We contrasted the learning economy to the standard no-learning economy with two shocks. Under no learning, a mean-preserving increase in the variance of TFP shocks increases long-run growth by raising precautionary savings. By contrast, a mean-preserving increase in the variance of investment shocks lowers growth because it raises the option value of waiting to invest, thereby reducing savings. Such a contrast does not arise in the learning economy: For the case of TFP shocks, the qualitative effects of learning work in the same direction as a permanent rise in variance. For shocks to the efficiency of investment, the qualitative effects can go in the opposite direction. in the learning economy growth rises whereas in the shock-variance-spreads economy the growth rate falls.

Endogeneity of growth reverses the implications for the risk-free rate: As in standard models, a rise in uncertainty raises precautionary savings, but because growth is endogenous, the rise in the growth rate is accompanied by a rise in the real interest rate, in contrast to what happens in an endowment economy.

When we compare the learning economy to one in which shock variances increase and equate the marginal distributions of the shocks at each date, the qualitative effects in the two cases do go the same way for TFP shocks, but quantitatively the effects are larger in the case of learning because expected lifetime utility varies more as beliefs respond to the shocks.

References

- [1] BANSAL, R. AND A. YARON, 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* 59, 1481-1509.

- [2] BAXTER, M., & JERMANN, U. J. (1997). The International Diversification Puzzle Is Worse Than You Think. *The American Economic Review*, 170-180.
- [3] BERNANKE, B., 1983. Irreversibility, uncertainty, and cyclical investment. *QJE* 98, 85-106.
- [4] BIANCHI, F., C. ILUT, AND M. SCHNEIDER. Uncertainty Shocks, Asset Supply and Pricing over the Business Cycle, *Review of Economic Studies*, 2018, 85(2), 810-854.
- [5] BIANCHI, F., AND L. MELOSI (2016). Modeling the Evolution of Expectations and Uncertainty in General Equilibrium. *International Economic Review* 57(2), 717-756.
- [6] BIANCHI, F., AND L. MELOSI (2018). Constrained Discretion and Central Bank Transparency. *Review of Economics and Statistics* 100(1), 187-202.
- [7] BLOOM, N. Fluctuations in Uncertainty. *J. Econ. Perspectives* 2014
- [8] CHAMLEY, C., & GALE, D. (1994). "Information revelation and strategic delay in a model of investment." *Econometrica*, 1065-1085.
- [9] COGLEY, T. 2009. Is the Market Price of Risk Infinite? *Economic Letters* 102(1), 13-16.
- [10] COLLIN-DUFRESNE, P., M. JOHANNES, AND L. LOCHSTOER. 2016. Parameter Learning in General Equilibrium: The Asset Pricing Implications. *American Economic Review* 106(3), 664-698.
- [11] DIXIT, A.K. AND R.S PINDYCK. 1994, *Investment Under Uncertainty*. Princeton University Press.
- [12] DOOB, J.L. Application of the Theory of Martingales. *Colloques Internationaux du Centre National de la Recherche Scientifique* 36, 1948.
- [13] EPSTEIN, L., & SCHNEIDER, M. (2008). Ambiguity, Information Quality, and Asset Pricing. *Journal of Finance*, 63(1), 197-228.

- [14] EPSTEIN, L.G. AND S.E. ZIN. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57(4), 937-969.
- [15] EVANS, G.W. AND S. HONKAPOHJA. *Learning and Expectations in Macroeconomics*, Princeton University Press: Princeton, 2001.
- [16] FISHER, J. D. (2006). The dynamic effects of neutral and investment-specific technology shocks. *Journal of Political Economy*, 114(3), 413-451.
- [17] FOERSTER, A., J. RUBIO-RAMIREZ, D.F. WAGGONER, AND T. ZHA. (2016) Perturbation Methods for Markov-Switching DSGE Models. *Quantitative Economics* 7, 637-669.
- [18] GEWEKE, J. 2001. A Note on Some Limitations of CRRA utility, *Economics Letters* 71(3): 341-345.
- [19] GILCHRIST, S. AND J. WILLIAMS. Investment, Capacity and Uncertainty: A Putty-Clay Approach, *Review of Economic Dynamics*, 2005.
- [20] GORDON N.J., SALMOND D.J. AND SMITH A.F.M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F* 140, 107-113.
- [21] GORDON, R.J. *The Measurement of Durable Goods Prices*. University of Chicago Press: Chicago, IL, 1990.
- [22] GORDON, R.J. *The Rise and Fall of American Growth*. Princeton University Press, 2016.
- [23] GOURIO, F. (2008). Time-series predictability in the disaster model. *Finance Research Letters*, 5(4), 191-203.
- [24] GREENWOOD, J., HERCOWITZ, Z., & KRUSELL, P. (2000). The role of investment-specific technological change in the business cycle. *European Economic Review*, 44(1), 91-115.
- [25] HALL, R.E. (1988). Intertemporal Substitution in Consumption. *Journal of Political Economy* 96(2), 339-57

- [26] HAMILTON, J. D. "A new approach to the economic analysis of nonstationary time series and the business cycle." *Econometrica* (1989): 357-384.
- [27] HANSEN, L., & SARGENT, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2), 60-66.
- [28] HAVRÁNEK, T. (2015). Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting. *Journal of the European Economic Association* 13(6), 1180-1204.
- [29] HOPENHAYN, H., M.E. MUNIAGURRIA. (1996). Policy variability and economic growth. *Review of Economic Studies*, 63 (4), 611-626.
- [30] HUFFMAN AND N. KIEFER. "Optimal Learning in a One-Sector Model." August 1994.
- [31] JONES, L. E., R. MANUELLI, H. SIU, AND E. STACCHETTI. "Fluctuations in convex models of endogenous growth, I: Growth effects." *Review of Economic Dynamics* 8, no. 4 (2005): 780-804.
- [32] JOVANOVIĆ, B. AND P. L. ROUSSEAU. (2005). "General Purpose Technologies." In P. Aghion and S.N. Durlauf, *Handbook of Economic Growth*, vol. 1B. Elsevier B.V., Amsterdam.
- [33] JOVANOVIĆ, B. AND P. L. ROUSSEAU. "Extensive and Intensive Investment over the Business Cycle." *Journal of Political Economy*. 2014.
- [34] JUSTINIANO, A., PRIMICERI, G. E., & TAMBALOTTI, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, 57(2), 132-145.
- [35] KENDRICK, J.W. *Productivity Trends in the United States*. National Bureau of Economic Research General Series no. 71. Princeton University Press, 1961.
- [36] KOULOVANTIANOS, C., L. MIRMAN AND M. SANTUGINI. "Optimal growth and uncertainty: Learning." *Journal of Economic Theory* 144 (2009) 280–295.
- [37] KREPS, D. Anticipated Utility and Dynamic Choice, 1997 Schwartz Lecture, in *Frontiers of Research in Economic Theory*, Edited by D.P. Jacobs, E. Kalai, and M. Kamien, Cambridge University Press, Cambridge, England, 1998.

- [38] KRUSSELL, P., L.E. OHANIAN, J.V. RIOS-RULL, G.L. VIOLANTE (2000). “Capital-Skill Complementarity: A Macroeconomic Analysis.” *Econometrica* 68, 1029-1053.
- [39] LEWELLEN, J., & SHANKEN, J. (2002). Learning, asset-pricing tests, and market efficiency. *Journal of finance*, 57(3), 1113-1145.
- [40] PÁSTOR, L., & VERONESI, P. (2009). Technological Revolutions and Stock Prices. *American Economic Review*, 99(4), 1451.
- [41] PIKETTY, T. *Capital in the Twenty First Century*, Harvard University Press: Cambridge, MA, 2014.
- [42] PINTUS, P & J. SUDA. (2018). Learning Financial Shocks and the Great Recession, *Review of Economic Dynamics*, forthcoming.
- [43] SARGENT, T. J. (1980). “Tobin’s q” and the rate of investment in general equilibrium. *Carnegie-Rochester Conference Series on Public Policy* 12, 107-154.
- [44] SIMS, C.A. 1982. Policy Analysis With Econometric Models, *Brookings Papers on Economic Activity*, 107-152.
- [45] STOKEY, N. “Wait-and-See: Investment Options under Policy Uncertainty.” *Review of Economic Dynamics* (2015).
- [46] VELDKAMP, L. L. (2005). “Slow boom, sudden crash.” *Journal of Economic theory*, 124(2), 230-257.
- [47] VERONESI, P. (2004). The peso problem hypothesis and stock market returns. *Journal of Economic Dynamics and Control*, 28(4), 707-725.
- [48] VISSING-JØRGENSEN, A. (2002). Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* 110(4), 825-853.
- [49] WACHTER, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility?. *The Journal of Finance*, 68(3), 987-1035.
- [50] WEIL, P. (1989). Non-expected Utility in Macroeconomics. *Quarterly Journal of Economics* 105(1), 29-42.

- [51] WHELAN, K. A Guide to US Chain Aggregated NIPA Data, *Review of Income and Wealth* 48, 2002, 217-233.
- [52] WRIGHT, S. “Measures Of Stock Market Value And Returns For The U.S. Nonfinancial Corporate Sector, 1900-2002.” *Review of Income and Wealth* 50, 2004.
- [53] Yogo, M. (2004). Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak. *Review of Economics and Statistics* 86(3), 797-810.

A Measuring the Output-Capital Ratio

In our model, TFP is the ratio of real output to the real capital stock. The left-hand column of figure 13 portrays raw data on real private-sector GDP and capital (upper panel) along with their ratio (lower panel).²³ As highlighted in the introduction, there is a structural break in the capital-output ratio in the 1940s.

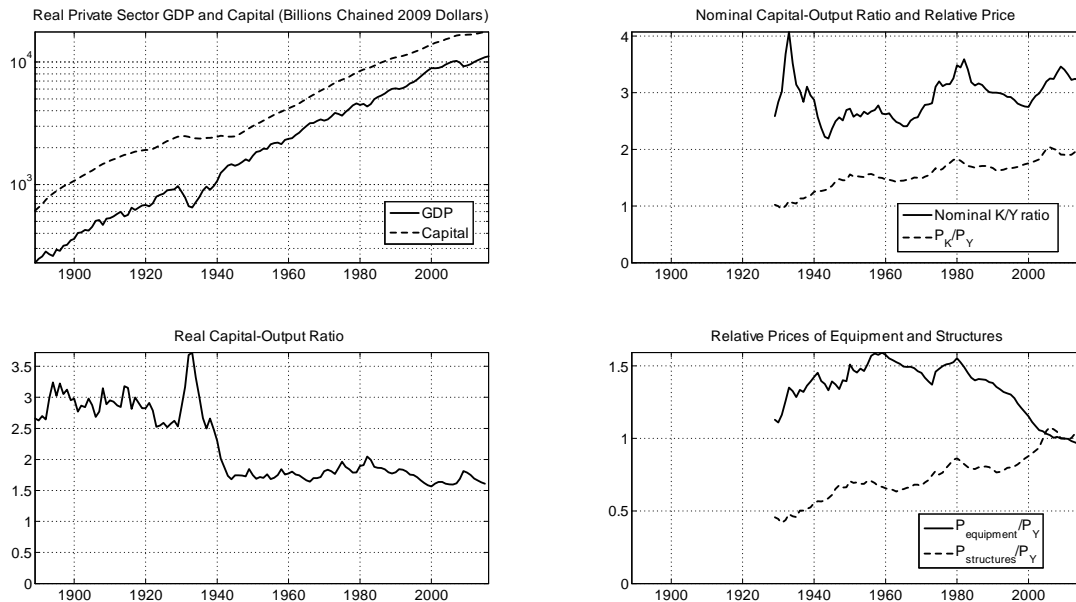


Figure 13: Output, Capital, and the Output-Capital Ratio

²³Data files and a program for calculating these measures are available on request.

The ratio of nominal capital to nominal output is shown in the upper right panel. Unlike the real ratio, the nominal ratio appears to be stationary, fluctuating around a value of 3.0. Since the nominal ratio is equal to the real ratio divided by the relative price of capital to output,

$$\frac{Y_{nt}}{K_{nt}} = \frac{z_t}{P_{Kt}/P_{Yt}},$$

this would be a valid measure of TFP if P_K/P_Y were always equal to 1. For instance, this would be true in a one-sector Ak model in which output could be costlessly transformed into capital and vice versa. Alas, the assumption that $P_K/P_Y = 1$ is not supported by the data. On the contrary, P_K/P_Y trends upward (see the upper right panel). The nominal ratio therefore confounds TFP with other shocks that move the relative price of capital to output.

Our measure of capital is constructed from data on equipment and structures. The lower right panel drills down by portraying the relative price of each component to output. Although the relative price of equipment to output declined sharply after 1980, the relative price of structures kept rising. Our measure of P_K is a Fisher ideal price index for the two components, and an upward trend emerges because the steadily rising price of structures more than counterbalanced the rise and fall in equipment prices. An upward trend in P_K/P_Y does not contradict evidence of a downward trend in P_E/P_Y .²⁴

²⁴Our relative-price data differ from those used in a number of influential papers in the literature. For instance, Jovanovic and Rousseau (2005) report a downward trend in the relative price of equipment for the entire period between 1885 and 1996. In contrast, our data show a rising relative price from 1925 to 1960 and a decline thereafter.

The difference is mainly due to how equipment prices are measured. Building on work by Krussell, et al. (2000), Jovanovic and Rousseau start with Gordon's (1990) data on quality-adjusted equipment prices for 1947-83, then extend it forward and backward in time by appending VAR forecasts and historical data (see Jovanovic and Rousseau 2005, footnote 9, p. 1196). We do not use Gordon's (1990) data. Instead, we follow Gordon (2016) by using NIPA data on chain-weighted price indices. These chain-weighted series incorporate many of Gordon's (1990) proposals, such as adjusting for quality by comparing like units in adjacent years and aggregating by computing superlative price indices. Coverage of equipment types is also broader; e.g. for 1967, Gordon (1990) reports that categories amounting to 23 percent of private domestic equipment was missing from his 1990 series. The new BEA data also cover a longer time span, 1925-2015 vs. 1947-93 for Gordon (1990), making forward and backward extensions unnecessary. Last but not least, Gordon (1990) collected no data on structures. Since our sources report comparable chain-weighted price indices for structures, we can construct a broader measure of aggregate capital.

A.1 Real private-sector GDP

The data are annual and cover the period 1889-2015. Two sources are spliced together at 1929. Data for the period 1889-1929 are taken from Kendrick (1961), while those for the period after 1929 were downloaded from the Bureau of Economic Analysis. Kendrick's series were rescaled so that they coincide with BEA measures at the splice date.

- 1889-1929: Real gross private domestic product was taken from Kendrick (1961), Appendix A, Table A-III, column 8.
- 1929-2015: Real GDP was taken from BEA Table 1.1.6, line 1. To adjust for government, nominal GDP and nominal government spending were downloaded from BEA Tables 1.1.5 (line 1) and 3.1 (line 26), respectively. The government share s_{gt} was measured as the ratio of nominal government spending to nominal GDP. Then real private GDP was calculated as

$$Y_{ps,t} = (1 - s_{gt})Y_t,$$

where Y_t is total GDP and $Y_{ps,t}$ is private-sector GDP.

A.2 Real Private-Sector Capital

The capital stock is defined as private fixed assets, structures plus equipment. Intellectual property was excluded because it was not available before 1925. Data for the period 1889-1925 come from Kendrick (1961), while those for the period 1925-2015 are from the Bureau of Economic Analysis. The two series are spliced in 1925, and Kendrick's series was again rescaled so that it coincides with the BEA measure at the splice date.

- 1889-1929: Real Structures and Equipment were taken from Kendrick, Appendix A, Table A-XVI, columns 7 and 9.
- 1925-2016: Index numbers for structures and equipment were downloaded from BEA Table 2.2, lines 2 and 41. The indices were converted to 2009 dollars by multiplying by nominal structures and equipment for that year. The latter can be found in BEA Table 2.1, lines 2 and 41, respectively. Following Whelan (2002), price and quantity indices for capital were computed by constructing Fisher ideal indices from the indices on equipment and structures.

B Proof of Proposition 1

Start with the planner's Bellman equation,

$$V(s_t, k_t) = \max_{C_t, X_t} \left[(1 - \beta)C_t^{1-\rho} + \beta \left((E_t[V(s_{t+1}, k_{t+1})^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}}) \right)^{\frac{1}{1-\rho}} \right],$$

then substitute the conjectured value function $V(s, k) = v(s)k$ along with the identities $C = ck$ and $k' = gk$:

$$\begin{aligned} v(s)k &= \max_{C, X} \left[(1 - \beta)(ck)^{1-\rho} + \beta \left((E[v(s')^{1-\gamma}(gk)^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}}) \right)^{\frac{1}{1-\rho}} \right], \quad (32) \\ &= \max_{C, X} \left[(1 - \beta)(ck)^{1-\rho} + \beta(gk)^{1-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}} \right], \\ &= \max_{C, X} \left[(1 - \beta)c^{1-\rho} + \beta g^{1-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}} \right] k. \end{aligned}$$

The aggregate resource constraint and law of motion for capital imply

$$c_t = z_t - q_t(g_t - 1 + \delta), \quad (33)$$

We use equation (33) to express the right-hand maximization problem in terms of growth,

$$v(s) \equiv \max_g \left[(1 - \beta)(z - q(g - 1 + \delta))^{1-\rho} + \beta g^{1-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}} \right]. \quad (34)$$

The first-order condition is

$$[(1 - \beta)(1 - \rho)(z - q(g - 1 + \delta))^{-\rho} q] = \beta(1 - \rho)g^{-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}}, \quad (35)$$

$$(1 - \beta)qc^{-\rho} = \beta g^{-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}}. \quad (36)$$

After substituting the conjectures for $c(s)$ and $g(s)$ in equations (17) and (16), we find

$$(1 - \beta)q \left(\frac{z + q(1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)} \right)^{-\rho} = \beta \left(1 - \delta + \frac{zq^{-\frac{1}{\rho}}w(s) - (1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)} \right)^{-\rho} (E[v(s')^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}}. \quad (37)$$

Multiplying both sides by $(1 + q^{1-\frac{1}{\rho}}w(s))^{-\rho}$,

$$\begin{aligned}
(1 - \beta)q(z + q(1 - \delta))^{-\rho} &= \beta \left[(1 - \delta)(1 + q^{1-\frac{1}{\rho}}w(s)) + zq^{-\frac{1}{\rho}}w(s) - (1 - \delta) \right]^{-\rho} \\
&\quad \times (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}, \\
&= \beta \left[(1 - \delta)q^{1-\frac{1}{\rho}}w(s) + zq^{-\frac{1}{\rho}}w(s) \right]^{-\rho} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}, \\
&= \beta w(s)^{-\rho} \left[(1 - \delta)q^{1-\frac{1}{\rho}} + zq^{-\frac{1}{\rho}} \right]^{-\rho} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}, \\
&= \beta w(s)^{-\rho} q [z + (1 - \delta)q]^{-\rho} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}. \tag{38}
\end{aligned}$$

Multiplying both sides by $w(s)^\rho$ and cancelling like terms implies

$$\begin{aligned}
w(s)^\rho &= \frac{\beta}{1 - \beta} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}, \tag{39} \\
w(s) &= \left(\frac{\beta}{1 - \beta} \right)^{\frac{1}{\rho}} \left[E \left((z' + q'(1 - \delta)) \left[(1 - \beta)(1 + q'^{1-\frac{1}{\rho}}w(s'))^\rho \right]^{\frac{1}{1-\rho}} \right)^{1-\gamma} \right]^{\frac{1-\rho}{\rho(1-\gamma)}},
\end{aligned}$$

This delivers a recursion for $w(s)$ that satisfies the FOC.

Now we need to verify that the conjectured solution satisfies the Bellman equation. Notice that

$$\begin{aligned}
v(s) &= \left[(1 - \beta)(z + q(1 - \delta)) \left(\frac{z + q(1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)} \right)^{-\rho} \right]^{\frac{1}{1-\rho}}, \\
&= [(1 - \beta)(z + q(1 - \delta))c^{-\rho}]^{\frac{1}{1-\rho}}. \tag{40}
\end{aligned}$$

If our guess holds, then (14) and (15) imply

$$\begin{aligned}
[(1 - \beta)(z + q(1 - \delta))c^{-\rho}]^{\frac{1}{1-\rho}} &= \left[(1 - \beta)c^{1-\rho} + \beta g^{1-\rho} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}, \\
(1 - \beta)(z + q(1 - \delta))c^{-\rho} &= (1 - \beta)c^{1-\rho} + \beta g^{1-\rho} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}} \\
\frac{z + q(1 - \delta)}{c} &= 1 + \left(\frac{g}{c} \right)^{1-\rho} \frac{\beta}{1 - \beta} (E[v(s')^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}} \\
\frac{z + q(1 - \delta)}{c} &= 1 + \left(\frac{g}{c} \right)^{1-\rho} w(s)^\rho \tag{41}
\end{aligned}$$

Plugging the conjecture for consumption in the left hand side yields

$$\begin{aligned}
1 + q^{1-\frac{1}{\rho}}w(s) &= 1 + \left(\frac{g}{c}\right)^{1-\rho} w(s)^\rho \\
q^{1-\frac{1}{\rho}}w(s) &= \left(\frac{g}{c}\right)^{1-\rho} w(s)^\rho \\
q^{1-\frac{1}{\rho}}w(s)^{1-\rho} &= \left(\frac{g}{c}\right)^{1-\rho} \\
q^{-\frac{1}{\rho}}w(s) &= \frac{g}{c}
\end{aligned} \tag{42}$$

Using the conjectured decision rules for c and g , the right-hand side can be written as

$$\begin{aligned}
\frac{g}{c} &= \left(\frac{z + q(1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)}\right)^{-1} \left(1 - \delta + \frac{zq^{-\frac{1}{\rho}}w(s) - (1 - \delta)}{1 + q^{1-\frac{1}{\rho}}w(s)}\right) \\
&= \frac{(1 - \delta)(1 + q^{1-\frac{1}{\rho}}w(s)) + zq^{-\frac{1}{\rho}}w(s) - (1 - \delta)}{z + q(1 - \delta)} \\
&= \frac{q^{1-\frac{1}{\rho}}(1 - \delta) + zq^{-\frac{1}{\rho}}w(s)}{z + q(1 - \delta)} w(s) \\
&= q^{-\frac{1}{\rho}}w(s)
\end{aligned} \tag{43}$$

From (42) and (43), it is clear that the conjectured value function and decision rules satisfy the Bellman equation.

C Alternating regimes when $\theta \in \{\theta_L, \theta_H\}$

Next we investigate a model with alternating regimes analogous to that of Hopenhayn and Muniagurria (1996, henceforth HM). A break occurs with probability λ . Conditional on a break, values of θ alternate with transition-probability matrix

$$\pi = \begin{array}{cc} & \begin{array}{cc} \theta_L & \theta_H \end{array} \\ \begin{array}{c} \theta_L \\ \theta_H \end{array} & \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \end{array} . \tag{44}$$

In other words, the timing of the shift is still random, but its character is deterministic. In the case of θ^q , this exercise is roughly that of HM, except that in their model a higher q is caused higher taxation (accompanied by a lump-sum redistribution) instead of higher investment cost.²⁵ As in HM, λ is an inverse measure of policy

²⁵HM study an Ak model that features policy shocks instead of investment-specific technology shocks. In their model, an investment subsidy is periodically put in place for a random length of

persistence. Higher variability then implies more frequent changes in consumption and investment.

Let us illustrate for the case where θ^z is fixed and θ^q oscillates as in (44). When we write “H” we mean the regime in which $\theta^q = \theta_H^q$. Beliefs are degenerate at the true θ^z, θ^q , and so (14) and Proposition 1 simplify as follows. For $I, J \in \{L, H\}$, (here We also assume CRRA preferences ($\rho = \gamma$), as in Sec. 3.

$$V^I(s, k) = \max_{C, X} \left\{ \frac{C^{1-\gamma}}{1-\gamma} + \beta \int (\lambda V^J(s', k') + (1-\lambda) V^I) dF^I(s') \right\} = v^I(s) k^{1-\gamma}.$$

where

$$v^I(s) = \frac{(z + (1-\delta)q)^{1-\gamma}}{1-\gamma} (1 + q^{1-\gamma-1} w^I(s))^\gamma$$

and where instead of (13), we have

$$w^I(s) = \left\{ \beta \int (z' + (1-\delta)q')^{1-\gamma} (1 + q'^{1-\gamma-1} (\lambda w^J(s') + (1-\lambda) w^I(s')))^\gamma dF^I(s') \right\}^{1/\gamma}$$

The optimal policies are

$$c^I(s) = \frac{z + (1-\delta)q}{1 + q^{1-\gamma-1} w^I(s)}, \quad \text{and} \quad g^I(s) = 1 - \delta + \frac{z q^{-1/\gamma} w^I(s) - (1-\delta)}{1 + q^{1-\gamma-1} w^I(s)}.$$

Then using again the parameters values in Table 3.1, Figure 14 plots the three averages

$$\frac{c^L(s_{\text{med.}}) + c^H(s_{\text{med.}})}{2}, \quad \frac{g^L(s_{\text{med.}}) + g^H(s_{\text{med.}})}{2} \quad \text{and} \quad \frac{w^L(s_{\text{med.}}) + w^H(s_{\text{med.}})}{2}$$

as functions of λ , evaluated at the median value of s .

When θ^q varies with θ^z fixed, more frequent policy shifts raise consumption and welfare because they allow agents to substitute investment from periods when investment is costly and into periods when it is cheap – see the bottom row of figure 14. This allows consumption to rise on average without an increase in its volatility. Welfare therefore goes up even though growth declines. On the other hand, regime changes that affect θ^z simply raise consumption volatility with no offsetting benefit from intertemporal substitution of investment – see the top row of figure 14.

As λ rises and as one moves from left to right along the horizontal axis in Fig. 14, future regimes become harder to predict as do future shock realizations. Although a time. This raises the growth rate similarly to how a fall in q raises growth in our model. We proceed, as HM do, under CRRA preferences ($\rho = \gamma$), as in Sec. 3.

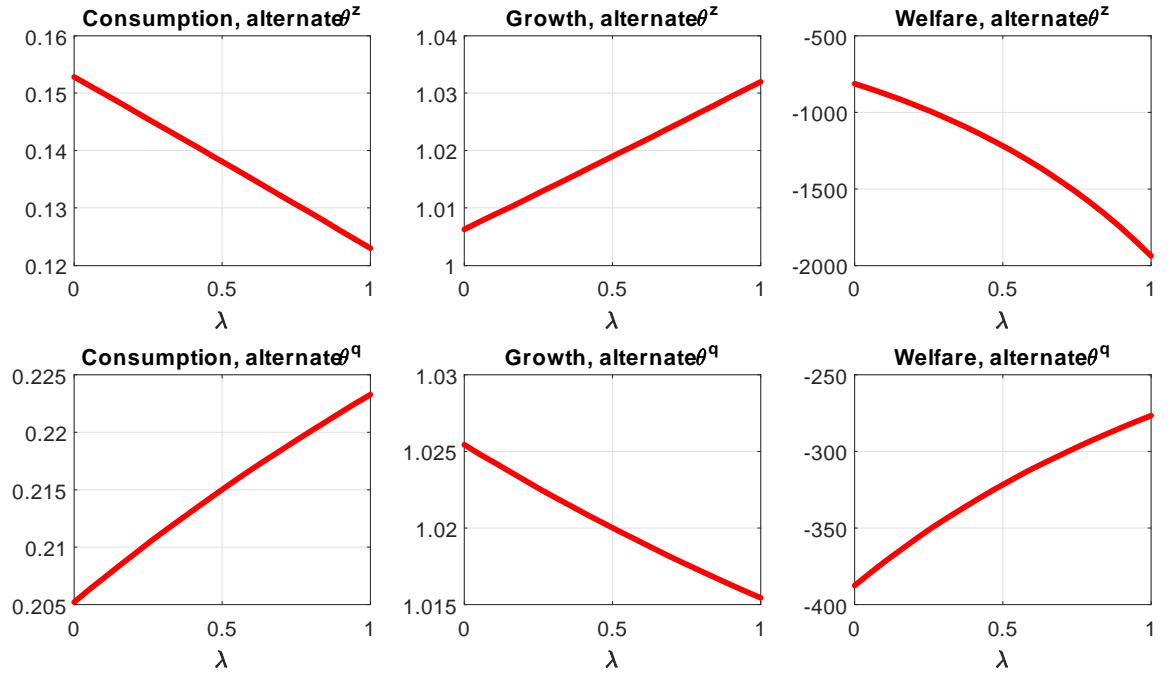


Figure 14: THE EFFECT OF λ WHEN θ ALTERNATES. TOP ROW: ONLY θ^z ALTERNATES, BOTTOM ROW: ONLY θ^q DOES.

rise in λ does not represent an MPS in the distributions of future shocks, these results are qualitatively similar to those in Sec. 3.1 that deal with MPSs for *given* and *known* values of θ . In Figure 2 shocks become less forecastable as we move *from right to left* on the horizontal axes in each of the four quadrants; consumption and growth respond in qualitatively the same way as in Fig. 14. But if the marginal distributions of the sequences of future shocks are the same for the two sets of cases, the agent is better off when regimes alternate because shocks are negatively autocorrelated and thus are easier to forecast in the alternating regime case.

D Solving the model with non-recurrent structural breaks and autoregressive shocks (section 4)

D.1 Agents use a pair of Kalman filters to learn about μ_t

The posterior $p(\mu_{t+1}, \mu_t | s^t, \theta)$ can be calculated recursively with a pair of Kalman filters. To simplify notation, define $x_t \equiv \ln a_t - \rho_a \ln a_{t-1}$. Because the history a^t

is observable and ρ_a is known, x_t is observable. Conditional on μ_{at} , x_t is normally distributed with mean μ_{at} and variance σ_a^2 .

Suppose a structural break occurs at date t .²⁶ Knowing that a new intercept is about to be drawn, agents adopt the unconditional distribution of m_t as their prior,

$$p(\mu_t|\chi_t = 1) = N(m, \sigma_m^2). \quad (45)$$

Next, agents observe x_t and update beliefs about μ_t . According to Bayes theorem,

$$p(\mu_t|x_t, \chi_t = 1, \theta) \propto p(x_t|\mu_t, \chi_t = 1, \theta)p(\mu_t|\chi_t = 1). \quad (46)$$

Because χ_t affects x_t only through μ_t , it carries no incremental information about x_t beyond that contained in μ_t . Hence the likelihood simplifies to

$$p(x_t|\mu_t, \chi_t = 1, \theta) = p(x_t|\mu_t, \theta). \quad (47)$$

Since the prior and likelihood are both Gaussian, this is a conjugate updating problem whose solution can be found by applying the Kalman filter,

$$p(\mu_t|x_t, \chi_t = 1, \theta) = N(\mu_{t|t}^m, 1/P_{t|t}^m). \quad (48)$$

A superscript m indicates that the mean and precision, $\mu_{t|t}^m$ and $P_{t|t}^m$, update the break-date prior (equation 45).

The predictive density at date t serves as the prior for $t+1$. To find this predictive density, we first condition on χ_{t+1} and then marginalize with respect to it. With probability λ , another structural break will occur at $t+1$, in which case μ_{t+1} will be distributed as

$$p(\mu_{t+1}|\chi_{t+1} = 1) = N(m, \sigma_m^2). \quad (49)$$

The new draw m_{t+1} is independent of the histories x^t or χ^t , making them irrelevant for prediction or updating.²⁷ With probability $1 - \lambda$, no break will occur, and $\mu_{t+1} = \mu_t$. For that scenario, the conditional predictive density is

$$p(\mu_{t+1}|x_t, \chi_t, \chi_{t+1} = 0) = N(\mu_{t+1|t}^k, 1/P_{t+1|t}^k), \quad (50)$$

²⁶The model is set up so that histories x^{t-1}, χ^{t-1} are cleared when there is a structural break, so this starting point involves no loss of generality.

²⁷This is why histories are cleared at break dates.

where superscript k indicates that the predictions are based on the output of Kalman updating at date t (equation 48). We marginalize with respect to χ_{t+1} by taking a probability weighted mixture,

$$\begin{aligned} p(\mu_{t+1}|x_t) &= \lambda p(\mu_{t+1}|\chi_{t+1} = 1) + (1 - \lambda)p(\mu_{t+1}|x_t, \chi_t, \chi_{t+1} = 0), \\ &= \lambda N(m, \sigma_m^2) + (1 - \lambda)N(\mu_{t+1|t}^k, 1/P_{t+1|t}^k). \end{aligned} \quad (51)$$

This serves as the prior for date $t + 1$.

At $t+1$, agents observe realizations of x_{t+1}, χ_{t+1} and update beliefs. The likelihood function is

$$\begin{aligned} p(x_{t+1}, \chi_{t+1}|\mu_{t+1}, \theta) &= p(x_{t+1}|\chi_{t+1}, \mu_{t+1}, \theta) \cdot p(\chi_{t+1}|\mu_{t+1}, \theta), \\ &= p(x_{t+1}|\mu_{t+1}, \theta) \cdot p(\chi_{t+1}|\lambda). \end{aligned} \quad (52)$$

In the second line, χ_{t+1} drops out of the first term because it conveys no incremental information about x_{t+1} beyond that contained in μ_{t+1} . In addition, μ_{t+1} drops out of the second term because χ_{t+1} is an exogenous Bernoulli random variable whose success probability does not depend on the intercept.

Multiplying the likelihood by the prior delivers the posterior kernel,

$$\begin{aligned} p(\mu_{t+1}|x_{t+1}, \chi_{t+1}, x_t, \chi_t, \theta) &\propto p(x_{t+1}|\mu_{t+1}, \theta) \cdot p(\chi_{t+1}|\lambda) \\ &\quad \cdot [\lambda p(\mu_{t+1}|\chi_{t+1} = 1) + (1 - \lambda)p(\mu_{t+1}|x_t, \chi_t, \chi_{t+1} = 0)]. \end{aligned} \quad (53)$$

Since $p(\chi_{t+1}|\lambda)$ does not depend on μ_{t+1} , this term can be absorbed into the normalizing constant, yielding

$$\begin{aligned} p(\mu_{t+1}|x_{t+1}, \chi_{t+1}, x_t, \chi_t, \theta) &\propto p(x_{t+1}|\mu_{t+1}, \theta) \\ &\quad \cdot [\lambda p(\mu_{t+1}|\chi_{t+1} = 1) + (1 - \lambda)p(\mu_{t+1}|x_t, \chi_t, \chi_{t+1} = 0)]. \end{aligned} \quad (54)$$

After substituting expressions for the priors, we find

$$\begin{aligned} p(\mu_{t+1}|x_{t+1}, \chi_{t+1}, x_t, \chi_t, \theta) &\propto p(x_{t+1}|\mu_{t+1}, \theta) \times N(m, \sigma_m^2) \quad \text{if } \chi_{t+1} = 1, \\ &\propto p(x_{t+1}|\mu_{t+1}, \theta) \times N(\mu_{t+1|t}^k, 1/P_{t+1|t}^k) \quad \text{if } \chi_{t+1} = 0. \end{aligned} \quad (55)$$

Both components involve a Gaussian prior and likelihood, so piecewise conjugacy is preserved. Hence the posterior is a mixture of normals, and each component can be updated via a Kalman filter.

In the event of a break, the prior is re-initialized to (45) and updated as at date t ,

$$p(\mu_{t+1}|x_{t+1}, \chi_{t+1} = 1, \theta) = N(\mu_{t+1|t+1}^m, 1/P_{t+1|t+1}^m). \quad (56)$$

Absent a break, the posterior is found via a standard Kalman update,

$$p(\mu_{t+1}|x^{t+1}, \chi_t, \chi_{t+1} = 0, \theta) = N(\mu_{t+1|t+1}^k, 1/P_{t+1|t+1}^k), \quad (57)$$

where x^{t+1} represents the history going back to the last break date.

The predictive density for $t+2$ has the same form as that for $t+1$. With probability λ , another structural break will occur at $t+2$, in which case μ_{t+2} will be distributed as

$$p(\mu_{t+2}|\chi_{t+2} = 1, \theta) = N(m, \sigma_m^2). \quad (58)$$

With probability $1 - \lambda$, no break will occur, and $\mu_{t+2} = \mu_{t+1}$. In that case, the conditional predictive density is

$$p(\mu_{t+2}|\chi_{t+2} = 0, x^{t+1}, \chi^{t+1}, \theta) = N(\mu_{t+2|t+1}^k, 1/P_{t+2|t+1}^k), \quad (59)$$

where x^{t+1}, χ^{t+1} are histories going back to the last break date. The predictive density is again a mixture of the two,

$$\begin{aligned} p(\mu_{t+2}|x^{t+1}, \chi^{t+1}, \theta) &= \lambda p(\mu_{t+2}|\chi_{t+2} = 1, \theta) + (1 - \lambda) p(\mu_{t+2}|\chi_{t+2} = 0, x^{t+1}, \chi^{t+1}, \theta), \\ &= \lambda N(m, \sigma_m^2) + (1 - \lambda) N(\mu_{t+2|t+1}^k, 1/P_{t+2|t+1}^k). \end{aligned} \quad (60)$$

This closes the loop. Updating and prediction at date $t+2$ have the same form as for $t+1$.

The updating densities are normal, and the predictive densities are mixtures of normal. Since conjugacy is preserved for each component of the mixture, the Kalman filter can be used to update conditional means and variances. Care must be taken about the prior being updated, but otherwise Kalman's logic goes through. When a break occurs, histories of x, χ are cleared, and the prior is re-initialized using the unconditional distribution. When no break occurs, histories of x going back to the last break date are retained, and updating is done with a Kalman filter.

D.2 How the integral equation is approximated

The state vector consists of the active shock $\ln a$ plus the posterior mean and precision, $\mu_{t|t}$ and $P_{t|t}$, that emerge from the Kalman filter. Because $P_{t|t}$ is a deterministic function of time since the last break, it can be subsumed in a time index for

$w(\cdot)$. We therefore look for a sequence of functions $w_t(\ln a_t, \mu_{t|t})$ for $t = 0, \dots, \infty$ that satisfy equation 13 and are consistent with Bayes updating and the law of motion for the shocks.

To simplify notation, define

$$f(s, w(s)) \equiv v(s)^{1-\gamma},$$

and write equation 13 as

$$w(s) = \left(\frac{\beta}{1-\beta} \right)^{\frac{1}{\rho}} [Ef(s', w(s'))|s]^{\frac{1-\rho}{\rho(1-\gamma)}}. \quad (61)$$

For a model with recurrent structural breaks, the expectation term has two branches,

$$\begin{aligned} Ef(s', w_{t+1}(s', \mu'))|s &= (1-\lambda) \int f \left(\left[s', w_{t+1}(s', \hat{b}_{nobreak}(s', \mu)) \right] \right) p_{nobreak}(s'|s) ds' \\ &+ \lambda \int f \left(\left[s', w_0(s', \hat{b}_{break}(s')) \right] \right) p_{break}(s'|s) ds'. \end{aligned} \quad (62)$$

A prime denotes next period's value, and t indexes time since the last break. $p_{break}(s'|s)$ and $p_{nobreak}(s'|s)$ represent the Bayesian predictive densities derived in the previous subsection, and $\hat{b}_{break}(s')$ and $\hat{b}_{nobreak}(s', \mu)$ signify that beliefs are updated according to Bayes rule. Substituting back into equation 13 yields

$$\begin{aligned} w_t(s, \mu) &= \left(\frac{\beta}{1-\beta} \right)^{\frac{1}{\rho}} \times \\ &\{ (1-\lambda) \int f \left(\left[s', w_{t+1}(s', \hat{b}_{nobreak}(s', \mu)) \right] \right) p_{nobreak}(s'|s) ds' \\ &+ \lambda \int f \left(\left[s', w_0(s', \hat{b}_{break}(s')) \right] \right) p_{break}(s'|s) ds' \}^{\frac{1-\rho}{\rho(1-\gamma)}} \end{aligned} \quad (63)$$

If λ were zero, the third line would vanish, and the first two would define a backward recursion for $w_t(\cdot)$. When $\lambda > 0$, the break date function $w_0(\cdot)$ affects the solution for $w_t(\cdot)$ for all non-break dates. The main challenge in solving equation (63) is the simultaneity between $w_t(\cdot)$ and $w_0(\cdot)$.

Our algorithm combines a backward recursion for the $w_t(\cdot)$ functions with a fixed point problem involving $w_0(\cdot)$. To find a good initial guess for the fixed-point problem, we adopt a method of successive approximation. Imagine a sequence of models indexed by λ_i , $i = 1, \dots, n$, starting from $\lambda_1 = 0$ and increasing until the desired value λ_n is reached. We first solve the $\lambda = 0$ model, iterating backward until $w_{0, \lambda_1}(\cdot)$ is

found. That solution is used to approximate $w_t(\cdot)$ for a model in which $\lambda_2 \approx 0$. We gradually raise λ to the desired value, approximating $w_0(\cdot)$ with the solution from the previous step. I.e., for $i = 2, \dots, n$, we solve

$$\begin{aligned}
w_{t,\lambda_i}(s, \mu) &= \left(\frac{\beta}{1-\beta} \right)^{\frac{1}{\rho}} \times & (64) \\
&\{ (1-\lambda_i) \int f \left(\left[s', w_{t+1,\lambda_i}(s', \hat{b}_{nobreak}(s', \mu)) \right] \right) p_{nobreak}(s'|s) ds' \\
&+ \lambda_i \int f \left(\left[s', w_{0,\lambda_{i-1}}(s', \hat{b}_{break}(s')) \right] \right) p_{break}(s'|s) ds' \}^{\frac{1-\rho}{\rho(1-\gamma)}}
\end{aligned}$$

This breaks the simultaneity between $w_t(\cdot)$ and $w_0(\cdot)$ and delivers a straightforward backward recursion for $w_t(\cdot)$. When the desired λ is reached, we iterate on

$$\begin{aligned}
w_{t,\lambda_n}(s, \mu) &= \left(\frac{\beta}{1-\beta} \right)^{\frac{1}{\rho}} \times & (65) \\
&\{ (1-\lambda_n) \int f \left(\left[s', w_{t+1,\lambda_n}(s', \hat{b}_{nobreak}(s', \mu)) \right] \right) p_{nobreak}(s'|s) ds' ds' \\
&+ \lambda_n \beta \int [z' + (1-\delta)q']^{1-\gamma} w_{0,\lambda_n}(s', \hat{b}_{break}(s')) p_{break}(s'|s) ds' \}^{\frac{1-\rho}{\rho(1-\gamma)}}.
\end{aligned}$$

until a fixed point for $w_0(\cdot)$ is found.

D.2.1 Solving the $\lambda = 0$ model

With $\lambda = 0$, one component of the mixture is deactivated, updating and prediction reduce to a conventional Kalman filtering problem, and equation (63) reduces to an infinite horizon backward recursion in $w_t(\cdot)$. We approximate its solution with that of a long but finite horizon problem.

The first step is to approximate the terminal value $w_T(\ln a_T, \mu_{T|T})$. Because the $\lambda = 0$ model involves decreasing gain learning, $\mu_{t|t}$ will eventually converge in probability to the true value. Agents don't know the true value, but they can calculate a rational expectations solution for $w(\cdot)$ for any value of μ_a in the support of the prior. Toward that end, we define grids for $\ln a_t$ and μ_a , set T equal to 1000 years, and shut down parameter uncertainty for year T and beyond by fixing $P_T^{-1} = 0$. For points on the grid, a fixed point for $w(\ln a, \mu_a)$ is found by iterating on the rational expectations version of the integral equation,

$$w(s, \mu) = \left(\frac{\beta}{1-\beta} \right)^{\frac{1}{\rho}} \left\{ \int f \left(\left[s', w(s', \mu) \right] \right) p(s'|s, \mu) ds' \right\}^{\frac{1-\rho}{\rho(1-\gamma)}}. \quad (66)$$

Values at ordinates between the nodes are approximated by fitting and interpolating a Chebychev polynomial. This delivers a family of rational expectations solutions indexed by μ_a that approximates the terminal solution $w_T(\ln a_T, \mu_{T|T})$.

This approximation initializes a backward recursion. For periods before T , parameter uncertainty is restored by updating $P_{t|t}$ via the Kalman filter, and $w_t(\cdot)$ is calculated by iterating backward on equation 63. Notice that the integrand on the second line involves $w_{t+1}(\ln a_{t+1}, \mu_{t+1|t+1})$. Next period's estimate $\mu_{t+1|t+1}$ is calculated using the Kalman filter, and $w_{t+1}(\cdot)$ is evaluated from the Chebychev approximation computed in the previous step of the backward recursion. The right-hand integral is then computed via quadrature. Values at ordinates between the nodes are again approximated by fitting and interpolating a Chebychev polynomial. Continuing back to date 0 delivers the desired sequence of $w_t(\cdot)$ functions.

D.2.2 Approximating equation (64)

Next we raise λ slightly and substitute $w_{0,\lambda=0}(\cdot)$ into equation (64). Since $w_{0,\lambda=0}(\cdot)$ is held constant, this equation also defines a backward recursion for $w_{t,\lambda_1}(\cdot)$, which is solved in essentially the same way as the $\lambda = 0$ model. We run through the sequence of models, gradually increasing λ , until an initial guess for $w_{0,\lambda_n}(\cdot)$, is obtained.

D.2.3 Solving the fixed point problem

We substitute the initial guess for $w_{0,\lambda_n}(\cdot)$ into the second line of equation (65), then compute Chebychev approximations to $w_{t,\lambda_n}(\cdot)$ by backward iteration, until a new guess for $w_{0,\lambda_n}(\cdot)$ is found. Then we substitute the new guess for $w_{0,\lambda_n}(\cdot)$ into the second line of equation (65) and re-solve the backward recursion, eventually finding yet another guess for $w_{0,\lambda_n}(\cdot)$. We continue until the maximum absolute difference between old and new guesses for $w_{0,\lambda_n}(\cdot)$ is negligible.

E Estimating parameters governing the shock processes

The shocks follows autoregressive processes with occasional shifts in the intercept,

$$\ln x_t = \mu_t + \rho_x \ln x_{t-1} + \sigma_x \varepsilon_{xt}, \quad (67)$$

where $x_t = z_t$ or q_t . A Bernoulli random variable χ_t governs whether a structural break occurs. With probability $1 - \lambda$, no break occurs ($\chi_t = 0$), and the intercept

remains unchanged. With probability λ , a break occurs ($\chi_t = 1$), and a new intercept is drawn from a normal distribution,

$$\begin{aligned}\mu_t &= \mu_{t-1} && \text{with pr } 1 - \lambda, \\ &= m_t \sim N(m, \sigma_m^2) && \text{with pr } \lambda.\end{aligned}\tag{68}$$

The random variables $\varepsilon_{xt}, \chi_t, m_t$ are mutually independent.

The parameters $\theta = [\lambda, m, \sigma_m, \rho_x, \sigma_x]$ are estimated from data on $\ln x_t$. A preliminary examination of the likelihood surface indicated that λ and σ_m are weakly identified in the frequentist sense. This is not surprising, since the sample might contain only a single structural break. Consequently, we add informative priors and find point estimates by maximizing the log posterior.

E.1 The likelihood function

As usual, the likelihood function can be factored as

$$p(\ln x^T | \theta) = p(\ln x_1 | \theta) \prod_{t=2}^T p(\ln x_t | \ln x^{t-1}, \theta).\tag{69}$$

Because the state-space representation is non-Gaussian, a particle filter is used to evaluate the likelihood function. Our particle filter is a straightforward modification of the Gordon, et al. (1993) bootstrap filter. It consists of three steps: predicting the state one period ahead, evaluating the period t component of the likelihood by Monte Carlo integration, and updating the date- t posterior for the hidden state.

E.1.1 State prediction

Suppose an evenly weighted sample from the $p(\mu_{t-1} | \ln x^{t-1}, \theta)$ is available at the end of period $t - 1$. With probability $1 - \lambda$, these particles will survive and move forward to date t . With probability λ , they will be extinguished and replaced with a draw $\mu_t = m_t \sim N(m, \sigma_m^2)$. The predictive distribution $p(\mu_t | \ln z^{t-1}, \theta)$ is a mixture of the two:

$$p(\mu_t | \ln x^{t-1}, \theta) = (1 - \lambda)p(\mu_t = \mu_{t-1} | \ln x^{t-1}, \theta) + \lambda p(\mu_t = m_t).\tag{70}$$

Because λ could be close to zero, we oversample the break distribution. Oversampling serves two purposes, injecting diversity into the particle cloud at each date and improving the accuracy of the monte carlo integration in step 2. Let

$$f(\mu_t | \ln x^{t-1}, \theta) = (1 - \tilde{\lambda})p(\mu_t = \mu_{t-1} | \ln x^{t-1}, \theta) + \tilde{\lambda}p(\mu_t = m_t)\tag{71}$$

represent a proposal density for $p(\mu_t | \ln x^{t-1}, \theta)$ with $\tilde{\lambda} > \lambda$. Our implementation sets $\tilde{\lambda} = 1/2$. Importance weights are adjusted below to compensate.

E.1.2 Evaluating the period t component of the likelihood

The period t component of the likelihood is

$$\begin{aligned} p(\ln x_t | \ln x^{t-1}, \theta) &= \int p(\ln x_t | \ln x^{t-1}, \theta, \mu_t) p(\mu_t | \ln x^{t-1}, \theta) d\mu_t, \\ &= \int p(\ln x_t | \ln x^{t-1}, \theta, \mu_t) \frac{p(\mu_t | \ln x^{t-1}, \theta)}{f(\mu_t | \ln x^{t-1}, \theta)} f(\mu_t | \ln x^{t-1}, \theta) d\mu_t. \end{aligned} \quad (72)$$

Armed with a proposal sample for μ_t from step 1, this can be approximated by

$$p(\ln x_t | \ln x^{t-1}, \theta) \approx N^{-1} \sum_{i=1}^N p(\ln x_t | \ln x^{t-1}, \theta, \mu_{it}) \frac{p(\mu_{it} | \ln x^{t-1}, \theta)}{f(\mu_{it} | \ln x^{t-1}, \theta)}. \quad (73)$$

where N is the number of particles,

$$\ln p(\ln x_t | \ln x^{t-1}, \theta, \mu_{it}) = -\ln \sigma_x - \frac{(\ln x_t - \mu_{it} - \rho_x \ln x_{t-1})^2}{2\sigma_x^2},$$

and

$$\begin{aligned} \frac{p(\mu_{it} | \ln x^{t-1}, \theta)}{f(\mu_{it} | \ln x^{t-1}, \theta)} &= \frac{\lambda}{\tilde{\lambda}} \quad \text{for draws from the break component,} \\ &= \frac{1 - \lambda}{1 - \tilde{\lambda}} \quad \text{for draws from the no-break component.} \end{aligned} \quad (74)$$

Our implementation sets $N = 2000$.

E.1.3 Updating the posterior for μ_t

The loop is closed by updating $p(\mu_t | \ln x^t, \theta)$. This is also done via importance sampling. The proposal density is the same as in step 1, and the target is the product of the conditional prior $p(\mu_{it} | \ln x^{t-1}, \theta)$ and the conditional likelihood $p(\ln x_t | \ln x^{t-1}, \theta, \mu_{it})$. Unnormalized importance weights are the ratio of the target to proposal,

$$\tilde{w}_{it} = p(\ln x_t | \ln x^{t-1}, \theta, \mu_{it}) \frac{p(\mu_{it} | \ln x^{t-1}, \theta)}{f(\mu_{it} | \ln x^{t-1}, \theta)}.$$

Sum and renormalize so that the weights sum to 1,

$$w_{it} = \frac{\tilde{w}_{it}}{\sum_i \tilde{w}_{it}}.$$

Then resample the proposal sample with weights w_{it} to get an evenly weighted sample from the date- t posterior $p(\mu_i | \ln x^t, \theta)$.

The particle filter then advances to period $t + 1$. Looping over t delivers a value for $p(\ln x^T | \theta)$.

E.2 Priors over TFP parameters

We assume prior independence, $p(\theta) = \prod_i p(\theta_i)$, and calibrate marginal priors for each parameter.

- $p(\lambda) = \text{beta}(8.98, 791)$. A beta prior is adopted so that $\lambda \in (0, 1)$. The prior mode was determined by counting trend breaks in the GDP data in Cogley (1990).²⁸ Casual inspection suggests roughly 1 trend break per 100 years in countries other than the US; hence the prior mode for λ was set to 0.01. Upper and lower values for a 95 percent prior confidence interval were set at values twice and half as large as the mode, respectively. A beta density with $\alpha = 8.98$ and $\beta = 791$ delivers the desired mode and confidence interval (see Gelman, et al., p. 476-77).
- $p(\rho_z) = \text{beta}(86.26, 13.74)$. A beta prior was also adopted to constrain ρ_z to the unit interval. The mode was found by simulating a conventional RBC representation for quarterly data,

$$\ln z_t = 0.95 \ln z_{t-1} + 0.006 \varepsilon_t. \quad (75)$$

Since the time period in our model is a year, we simulate a long quarterly series, transform into annual data by summing across quarters, and then estimate an $AR(1)$ model for the annual data. The result implies a prior mode of $\rho_z = 0.87$. The prior 95 percent confidence interval was then set to $\rho_z \in [0.79, 0.925]$. A beta density with $\alpha = 86.26$ and $\beta = 13.74$ delivers the desired mode and confidence interval.

- $p(\sigma_z) = IG_1(25, 0.0387)$. The time aggregation in the last bullet point delivered an estimate of $\sigma_z = 0.0368$ for time-aggregated annual data. We also target a prior 95 percent confidence set of $\sigma_z \in [0.03, 0.055]$. An IG_1 density with shape coefficient 25 and scale parameter 0.0387 delivers the desired mode and approximates the confidence set for σ_z .

²⁸A mean shift in TFP generates a trend break in an AK growth model.

- $p(m) = N(-0.0901, 0.0033)$. Given the other prior parameters, $p(m)$ was calibrated so that the center of the unconditional distribution for structural breaks is near $K/Y = 2$.
- $p(\sigma_m) = IG_1(30, 0.0158)$. Similarly, the prior for σ_m is calibrated so that the structural break distribution concentrates on $K/Y \in (1, 3)$ with high probability.

The priors on (m, σ_m, ρ_z) jointly imply the following structural break distribution for the capital-output ratio. The prior allows both growth miracles and growth disasters after a structural break. It is also asymmetric, skewed toward growth disasters.

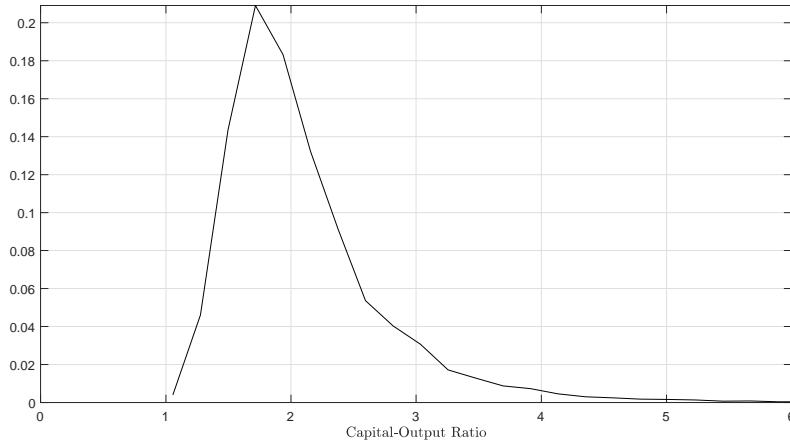


Figure 15: Implied prior for structural breaks in the capital-output ratio

E.3 Priors over investment shock parameters

We again assume prior independence, $p(\theta) = \prod_i p(\theta_i)$ and calibrate marginal priors for each parameter.

- Lacking better information, the prior for λ is the same as above, $p(\lambda) = \text{beta}(8.98, 791)$. This puts the mode at $\lambda = 0.01$ with a 95 percent prior credible set of $\lambda \in (0.005, 0.02)$.

- $p(\rho_q) = \text{beta}(5, 5)$. A beta prior constrains ρ_q to the unit interval. The hyperparameters were chosen so that the mode is 0.5 and a prior 95 percent credible set is $\rho_q \in (0.2, 0.8)$. This prior is weakly informative, and ρ_q is relatively well identified by the data, so this component of the prior is less influential than some of the others.
- $p(\sigma_q) = IG_1(5, 0.24)$. This prior is also weakly informative, with a prior mode of 0.2 and a 95 percent credible set of $\sigma_q \in (0.13, 0.5)$. Our main concern was to allow transient innovations to be highly volatile, reflecting the high volatility of market valuations.
- $p(m) = N(0, 0.025)$. Given the other prior parameters, $p(m)$ was calibrated so that the center of the unconditional distribution for m is close to 0.
- $p(\sigma_m) = IG_1(10, 0.0229)$. Last but not least, the prior for σ_m is calibrated so that, after a structural break, the new anticipated-utility long-horizon forecast for q lies between 0.5 and 1.5 with probability 0.99.

The priors on (m, σ_m, ρ_q) jointly imply the following structural break distribution for conditional long-horizon forecasts of q .

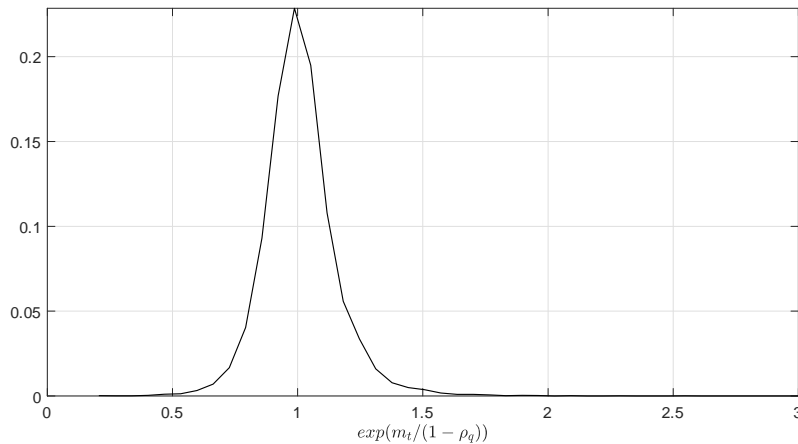


Figure 16: Implied prior on conditional long-horizon forecasts for q